Université catholique de Louvain
Faculté des sciences économiques, sociales, politiques et de communication
Economics School of Louvain
Louvain Institute on Data Analysis and Modeling
Institut de Recherches Economiques et Sociales

Human Capital and Knowledge Networks

Market Forces, Scientific Academies, and the Spread of Ideas in Early Modern Europe

---

DOCTORAL DISSERTATION PRESENTED BY

Chiara Zanardello

IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREES OF
DOCTORAT EN SCIENCES ÉCONOMIQUES ET DE GESTION

Thesis Committee:

Prof. Luca Pensieroso, *president*	UCLouvain
Prof. David de la Croix, *promotor and secretary*	UCLouvain
Prof. Joseph F. Gomes	UCLouvain
Prof. Ralf R. Meisenzahl	Chicago FED
Prof. Paula Gobbi	ULB

Louvain-la-Neuve, Belgium
September 2, 2025

Dedicato ai miei angeli custodi,

Antonietta e Guerrino
Agnese e Giuseppe

Acknowledgements

This thesis is the realization of a long-held dream, yet I could never have imagined the incredible journey that would lead me to this point. Every person, every place, and every experience has become an indispensable part—not only of this work, but of who I am today. So, thank you!

Thanks to my supervisor, David. You believed in me since my Master's thesis, and your competence, precision, and generosity guided me throughout the last five years. I will never be grateful enough for teaching me not only how to conduct meaningful and independent research, but also for showing me that kindness goes a long way. Without your optimism, these years would have been much less inspiring and fun.

Thanks to my committee and jury members. Luca, thanks for all the lessons you taught me during these years: to push my reasoning beyond the technical details of empirical applications, to always tell an engaging story with my research, and to always get fruit at American conferences. Joseph, thanks for all your valuable comments, and also for all the funny and relaxed corridor chats.

Paula, thanks for all the advice, energy, and time you spent restructuring my presentations and papers. Ralf, thanks for your insightful contributions, which pushed my papers to another level. It was all extremely helpful.

I would also like to thank the faculty and friends at Northwestern University, especially Joel Mokyr and Walker Hanlon, as well as Myera, Franco, Pedro, and Johanna, for welcoming me and making me feel part of the team.

A special thanks to Mara and Fabio for the academic discussions and for making my first stay in Providence much easier. Thank you also to Matthew Curtis, whose comments and generosity not only helped shape my job market paper but also supported me through that particularly challenging period.

Thanks to IRES. I had so much fun during my years there, and I want to thank my fellow PhD students for that. Giulia, I am so glad we entered this journey together after sharing some stressful but exciting months during our Research Master. I can always count on you for a chat, a technical clarification, an apéro, or a place to stay—thank you! I would like to thank Filippo and Rossana too; I am glad you decided to come on board the econ history team—it's much more fun with you around! Thanks also to Mathilde for sticking around IRES and always being available for a much-needed drink at the end of a tough week. Thanks to Gaia for sharing coffee breaks, research advice, post-work drinks, and the office D.205.

I would also like to thank all the other PhD friends I met during these years: Alessandro, Angel, Anna, Arnaud, Arno, Charles, Daniele, Diego, Ella, Esme, Femke, Florian, François, Francesco, Gianmarco, Guillermo, Hakimeh, Janie, Kam, Keiti, Laura, Leo, Luigi, Margaux, Marion, Mathilde, Mauricio, Nathan, Natalia, Quentin, Riccardo, Robin, Rosa, Silvia, Vincent, Yannik.

Thanks to my family. Grazie ai miei genitori, Francesco e Cristina, per aver creduto in me fin da subito, per avermi lasciato andare ovunque volessi, e per tenere sempre la porta di casa aperta per tutte le volte in cui sento il bisogno di tornare. Capisco tutti i sacrifici che avete fatto per farmi arrivare a questo punto e ve ne sarò per sempre grata. Vi potrò sembrare lontana quando viaggio in giro per il mondo, ma ad ogni passo che faccio capisco sempre di più quanto sono fortunata ad avere una famiglia così alle mie spalle. Grazie anche ai miei fratelli. Grazie a te, Fabio, per farci capire che c'è sempre qualcosa in più che possiamo fare per migliorare questo mondo. Grazie a Davide per avermi spinto a superare le mie paure e a partire per l'Erasmus… te l'ho già detto molte volte, ma senza quella spinta non sarei dove sono ora. Grazie!

Vorrei ringraziare anche tutto il resto della famiglia, zii e cugini. In particolare, zia Tina e zio Gianni, per avermi accolto e ospitato durante i miei anni di triennale, permettendomi di concentrarmi a pieno nei miei studi. Grazie anche a Marta, Roberto, Cate e Vicky: senza di voi non avrei avuto accesso a tutte le possibilità che l'università di Milano mi ha messo a disposizione. Senza quei mesi da voi non ci sarebbe stata nessuna doppia laurea e non avrei (forse mai) scoperto l'esistenza di Louvain-la-Neuve.

Thank you also to my new Belgian family. Merci à Géraldine et Madi de m'avoir toujours accueillie avec un grand sourire et de délicieux gâteaux. Merci à Christophe, Gérard et Nicole pour votre soutien et votre engagement dans nos vies. Grâce à vous, la Belgique est devenue comme une seconde maison pour moi.

Thanks to my friends, scattered around the world—from Italy, to Australia, passing via Belgium and the US. Thanks for coloring my days from wherever you are. A special thanks to Ceci and Ines: you have been there since my first Belgian days, and without you they would have been much more gray.

Last, but not least, thanks to Antoine. You made every step we took together much lighter and brighter. I am curious to see where the next ones will take us.

Contents

Acknowledgements **iii**

0 General introduction **1**
Bibliography 3

1 Market forces in Italian Academia today (and yesterday) **5**
1.1 Introduction 6
1.2 Data sampling 9
1.3 Methodology 20
1.4 Comparison between the present and the past 28
1.5 Conclusions 33
1.6 Bibliography 34
1.A Additional tables 38
1.B Standard Logit Model - technical details 42
1.C Age analysis 44
1.D Multiple Affiliations analyses 46
1.E Gender analysis 51
1.F Nested Logit Model - Technical Details 53
1.G Private/Public universities - further insights 58
1.H Role of sorting 63
1.I Main regressions with the threshold at 5 scholars 64

2 Early Modern Academies, Universities and Growth **69**
2.1 Introduction 70
2.2 Historical and institutional context 74
2.3 Data 78
2.4 Empirical strategy 84
2.5 Results 86
2.6 Quality of Universities 92
2.7 Additional analyses 94
2.8 Conclusions 99

2.9 Bibliography 101
2.A Additional details on historical context 107
2.B Descriptive Statistics 128
2.C Additional results 137
2.D Alternative DID estimators 147
2.E Robustness checks 152

3 Flora, Cosmos, Salvatio **175**
3.1 Introduction 176
3.2 Data and Methodology 182
3.3 Structural estimation 195
3.4 Further empirical assessments 204
3.5 Counterfactual experiments 213
3.6 Conclusions 224
3.7 Bibliography 225
3.A Descriptive Statistics 230
3.B Contextual information 232
3.C Empirical Assessments 236
3.D Tests for the proportionality of Hazard Functions 241
3.E Example of scholastic reasoning 242
3.F Augmenting Rubin's data 242
3.G Scholasticism: Additional Results 244
3.H Additional Results with Practical Surgery 246
3.I Placebo Inventors of Botanical Realism 248
3.J Placebo networks and the role of the Jesuits 250

General introduction

Why Human Capital?

This thesis explores the role that high-level knowledge and skills have played in shaping Europe's economic growth over the past millennium. I focus on academic scholars—a small segment of the population that has, arguably, preserved the highest levels of human capital from 1000 CE to the present. While today's economy may be more strongly influenced by other knowledge elites, such as professional engineers and skilled entrepreneurs, these groups emerged only in the 18th century (Hanlon, 2022), following transformative episodes like the Enlightenment and the Scientific Revolution.

Adopting a long-term historical perspective is essential for grasping the complexity of this topic. The literature broadly agrees that advanced skills are crucial—if not indispensable—for the prosperity of modern developed societies, yet the role of high-quality knowledge before the Industrial Revolution remains contested. By tracing investments in human capital and advanced education over centuries, I am able to shed new light on their long-run effects and evolution. This thesis contributes to the debate through new data and innovative empirical strategies, with each of the three chapters addressing a distinct aspect of the question.

Market Forces. The first chapter, "Market Forces in Italian Academies today (and yesterday)," focuses on the mobility of professors currently active in Italian universities to investigate the factors influencing their choices of where to teach. I construct a novel database containing individual-level information on their place of birth, place of undergraduate education, current place(s) of employment, age, and gender. Additionally, I develop a composite indicator of individual quality based on citation metrics.

I model professors' location choices as a function of distance, individual and institutional quality, revealing insightful trends. By comparing the results of this modern database with the historical dataset from other projects and detailed in de la Croix (2021), I offer a fresh perspective on Italy's contemporary academic market, while also providing relevant policy recommendations. Initially, I expected the Italian academic market to be static, conservative, and lacking in meritocracy. However, the evidence reveals some encouraging signs that suggest the potential for a virtuous cycle to emerge. In particular, my findings on the historical evolution of the sorting effect—where higher-quality scholars place greater value on university quality—show that during the Renaissance, Italian universities were highly successful in attracting top scholars. This strength, however, declined between 1530 and 1800, as new universities emerged and other institutions began favoring personal networks over merit. Encouragingly, by 2020, signs of renewed sorting have reappeared: recent recruitment reforms may be enhancing the influence of university quality while limiting nepotism and favoritism.

Nevertheless, future research should broaden the scope to encompass all of Europe, allowing for the integration of international mobility patterns. It should also extend the temporal range to examine the effects of recent policies, such as the 2019 tax incentives aimed at mitigating the Italian brain drain.

Scientific Academies. In the second chapter, “Early Modern Academies, Universities, and Growth," I examine the role of scientific academies in the economic development of European cities before 1800. As with the previous chapter, I manually collected individual-level data on members of scientific academies across Europe between 1500 and 1800. I also compiled aggregate information to investigate the origins of these academies, how they differed from traditional institutions such as universities, and whether complementarities between the two can be identified.

The first part of my findings suggests that scientific academies channeled talents and resources toward economically productive activities, while literary academies tended to divert human capital toward lower-/no-return endavors. The second part of the project reveals complementarities between academies and universities that extend beyond direct urban growth. I consider this the most original contribution of the chapter: I provide empirical evidence that scientific academies pushed co-located universities to improve their quality, undertake reforms, and pursue innovation.

To the best of my knowledge, these are the first quantitative results to support the historical, qualitative evidence of synergies between these two distinct types of educational institutions (Applebaum, 2000; McClellan, 1985)—synergies that contributed to the emergence of more professionally oriented universities as we know them today. While these findings represent a significant step forward, future projects could further explore the origins of the modern educational system by deepening this line of research.

Spread of Ideas and Knowledge Networks. The third chapter completes the picture by examining the role of networks—formed by the institutions studied in the previous chapters, universities and academies—in the dissemination of ideas. Entitled “Flora, Cosmos, Salvatio: Pre-modern Academic Institutions and the Spread of Ideas," this chapter is a joint project with David de la Croix and Rossana Scebba. Drawing on data collected through the ERC project (de la Croix, 2021) and my work from the second chapter, we investigate the network formed by scholars’ institutional affiliations. In our framework, two scholars are linked if they belong to the same institution, in the same year, and work within the same broadly defined field.

This setting provides a valuable opportunity to systematically analyze the diffusion of knowledge across institutions. Unlike citation or co-authorship networks, it does not rely on compliance, and unlike correspondence networks, it benefits from more complete and structured data with identifiable gaps. Most importantly, it allows us to highlight the specific role each institution plays in the dissemination of knowledge—our central aim.

Using this affiliation network, we identify an “inventor”—the first scholar to develop a groundbreaking idea—and simulate how that idea might spread across the network. This approach enables us to measure varying levels of exposure to the idea over time and space. While this project focuses on a network of individuals, future research could take a step further: by applying advanced text-mining algorithms, it may be possible to construct a network of ideas and link it to our scholar network. The implications would be profound. First, it would allow for the study of historical innovations in a period predating the patent system. Second, it could offer deeper insights into the historical relationship between fundamental research and practical applications (Brooks, 1994).

Bibliography

Applebaum, W. (2000). *Encyclopedia of the scientific revolution: from Copernicus to Newton*. Routledge.

Brooks, H. (1994). The relationship between science and technology [Special Issue in Honor of Nathan Rosenberg]. *Research Policy*, *23*(5), 477–486.

de la Croix, D. (2021). Scholars and Literati in European Academia before 1800. *Repertorium Eruditorum Totius Europae*, *5*, 35–41.

Hanlon, W. W. (2022). *The rise of the engineer: Inventing the professional inventor during the Industrial Revolution* (tech. rep.). National Bureau of Economic Research.

McClellan, J. E. (1985). *Science Reorganized: Scientific Societies in the Eighteenth Century*. Columbia University Press.

Chapter 1

Market forces in Italian Academia today (and yesterday)*

This paper investigates the operation of the academic market in Italy, mapping current scholars' location choices. I build a new dataset of current professors, associating each scholar with a composite indicator of their quality. The analysis includes the quality of the university and the features of the city where the institution is located. I estimate the strength of different factors: gravity (distance), agglomeration (scholars are attracted to higher quality universities), selection (better scholars travel longer distances), and sorting (the better the scholar, the more the quality of universities is weighted). I find that all of these factors have an effect, and do not vary according to scholars' gender. I find a greater expected utility for scholars in choosing private universities over public ones, through a consistent nesting procedure. Comparing these forces to historical trends in Italian academia, the sorting effect delineates a new momentum for the current academic market in Italy.

*This article is published and the material is taken from: Zanardello, C. Market forces in Italian academia today (and yesterday). Scientometrics 128, 651–698 (2023). https://doi.org/10.1007/s11192-022-04579-0.
I thank David de la Croix, Massimiliano Bratti, Daniele Checchi, and Fabrizio Iacone, who have contributed since the early draft of the project. I also thank all the scholars who replied to the survey. Simone Bertoli and Kenneth Train provided useful technical advice.
The first draft of this research (Zanardello, 2022) can be found as a working paper at http://hdl.handle.net/2078.1/258573.

1.1 Introduction

Knowledge and knowledge mobility have enormous social impact. The location choice of scholars, who are in the upper tail of the human capital distribution, has a measurable effect on knowledge formation, productivity and innovation. Bahar and Rapoport (2018) show that the location choice of high-skilled migrants draws the knowledge diffusion path, which has relevant effects on productivity. Location and birthplaces analyses can also disentangle economic competitiveness and complexity. When there is skill-complementarity, more diverse birthplaces result in more complex economic systems and higher economic growth (Alesina, Harnoss, & Rapoport, 2016; Bahar, Rapoport, & Turati, 2019; Docquier et al., 2020). The connection between human capital and modern economic growth is also confirmed in theoretical and empirical results. The essential role of education for adaptating to changing contexts and for driving modernization was first claimed by Nelson and Phelps (1966). Lucas Jr (1988) modeled the positive spillover effects of education, and more recently many empirical studies have presented clear evidence about the essential role of human capital and its institutions in improving current societies (Agasisti, Barra, & Zotti, 2019; Barra & Zotti, 2017, 2018; Barro, 1991, 2001; Beine, Docquier, & Rapoport, 2001; Cohen & Soto, 2007; Cottini, Ghinetti, & Moriconi, 2019; Hanushek & Woessmann, 2008).

Notwithstanding its fundamental role in socio-economic progress, investments in knowledge and educational institutions are not linear over time. There is a non-steady process of human capital accumulation, with periods of growth, decline, and recovery that may seed a new cycle of expansion (Artige, Camacho, & de La Croix, 2004). Italy is an interesting example of this fluctuating process: it dominated intellectual activities until the late Renaissance, but its contemporary cultural and educational system has weak incentives (Checchi, De Fraja, & Verzillo, 2021). Italian universities outperformed learning institutions in other countries from the 14th century to the first part of the 16th century (de la Croix et al., 2023), but only three universities are counted among the 200 top institutions in the 2022 QS World University Rankings[2].

I study the mobility of modern day Italian academics, and how they choose where to develop their career, to learn more about the modern university system of the peninsula. I take several factors into account: distance is increasing the cost of travelling, but it might be offset by the skills and knowledge acquired by the scholar and by the prestige of the university. I estimate professors' location choice as a function of distance and quality, given the location of universities. I

[2] Polytechnic of Milan - 142nd, University of Bologna - 166th and Sapienza University of Rome - 171st.
2022 QS World University Rankings are at: https://www.topuniversities.com/university-rankings/world-university-rankings/2022.

map the current academic market with its professors' human capital and compare this to previous eras of Italian academia. By comparing how scholars are moving nowadays to how they moved in the past, I locate Italy in the fluctuating cycle, estimate the path of Italian academia, and map out future directions.

For my research I have built a new dataset of contemporary Italian professors in the economic field, capturing information about their origins and their individual quality. To collect information about the birthplaces of live persons I had to secure privacy authorizations, which could have hindered data collection. To overcome this missing data issue, I used a more accessible proxy: the location of professors' lowest level of education. Once I had a value for the birth and/or education location, I used a Principal Component Analysis (PCA) to build a composite indicator of individual quality out of eight bibliometric indexes.

I use a Random Utility Model (RUM), and specifically, a multinomial logistic regression to compare scholars' utility of living in a region other than their birthplace (Beine, Bertoli, & Fernández-Huertas Moraga, 2016; Bertoli & Moraga, 2013; Bertoli & Rapoport, 2015; Ortega & Peri, 2013). I limit the analysis to choices made within the academic world, given the impossibility to consider the choice faced by academics when they decide whether to become a professor or to follow other career paths. I use the approach developed by de la Croix et al. (2023), who study the European Academic Market from 1000 CE to 1800 CE, to compare past and present outcomes in academia. My main estimations rely on information about geographical distance, individual quality of current professors (*human capital* hereafter), and aggregate quality of Italian universities (*notability* hereafter). (1) *Agglomeration* investigates whether Italian scholars are attracted by universities with higher notability, (2) *positive sorting* tests whether scholars with higher human capital weigh the notability of universities higher than do professors with lower individual quality, through the interaction term between human capital and notability, and (3) *positive selection* questions whether scholars with higher human capital move further, utilising the interaction between human capital and distance. There is an extensive literature showing that better-educated individuals are the most mobile portion of the population, with their higher growth perspectives giving them stronger incentives to move (Beine, Bierlaire, & Docquier, 2021; Docquier & Marfouk, 2006; Faggian & McCann, 2009; Handler, 2018; Schiller & Cordes, 2016; Zhao et al., 2021). Grogger and Hanson (2011) showed precisely that highly educated people are more likely to move (positive selection) and that these highly-specialised migrants choose destinations that compensate knowledge better (positive sorting).
My hypotheses assume the presence of strong complementarity in knowledge and skills (Easterly, 2001), which leads to positive assortative matching: working with better scholars (i.e., with higher human capital) would increase the marginal gain of each professor, and this increase is greater for better scholars than for

academics with lower human capital (Kremer, 1993). High complementarity also explains why more notable universities, populated with better scholars, attract more and more high-quality professors which may initiates a virtuous cycle of human capital accumulation.

I estimate these effects for Italian academia, and find that the standard distance effect is negative and has a magnitude in line with migration literature (Beine, Docquier, & Özden, 2011). To study *agglomeration* I include attractive features of the city in which the university is located (size and wealth), in addition to notability. The latter is positive but not significant, signalling that the relevance Italian universities' quality can be improved by public policy to increase its attractive power for contemporaneous scholars. Agglomeration is instead driven by the disposable household income in the city (i.e., city wealth). However, the estimator for the size of the city is negative, implying a dispersion effect – although with a lower magnitude than agglomeration. This finding is crucial for understanding mobility patterns and policy directions: it is essential to attract high-skilled people to create a dynamic context and generate positive spillovers for society, which may lead the country into the virtuous part of the cycle (Grogger & Hanson, 2015; Kerr et al., 2016, 2017; Stephan & Levin, 2001). I also find evidence of positive selection and positive sorting. Indeed, *positive selection* (interaction between distance and individual quality) is a solid result, which confirms that the higher the individual quality, the stronger the incentives for the scholar to travel to progressively better destinations. *Positive sorting* (interaction between human capital and notability) has a weaker significance level than selection. The weakness of sorting is due to the structure of Italian higher education, which may be still influenced by the traditional seniority-based system (Capano, 2008; MacLeod & Urquiola, 2021; Rebora & Turri, 2008). Reforms to increase the autonomy of universities, like the decentralization of the recruitment process of university professors,[3] may have led to the greater importance of local excellence for mobility decisions. However, such reforms are too recent to be in the current project, and so sorting only reaches a low level of significance. The seniority-based system may explain why Italian universities lost their leading position: there is evidence of a highly significant positive sorting only until 1526, which fades towards 1800. The sorting effect only regains power in the sample of present-day scholars, but the significance of current sorting is not as strong as at the birth of Italian universities. This is probably due to the very recent academic reforms. Either way, these results are key to understanding Italy's current position in the cycle, where it seems to be new momentum for Italian universities. In addition to the main regressions, I test for gender differences and find no significant outcomes. Men and women have similar patterns of mobility in Italy, but women represent only 30% of the sample.

[3]DPR n. 390/1998, law n. 210/1998, and law n. 240/2010, among the others.

I do find important differences between public and private universities. A variant of the standard logit model shows a greater expected utility for scholars in choosing private universities over public ones. This bolsters the argument in favour of a more autonomous, excellence-driven academic apparatus.

My analysis contributes to the migration and knowledge-based mobility literature. To the best of my knowledge, much of this literature deals with more general samples of high-educated/high-skilled people (Beine, Docquier, & Özden, 2011; Docquier & Marfouk, 2006; Grogger & Hanson, 2011; Handler, 2018; Kerr et al., 2016, 2017). Only a few articles investigate the mobility of academics or scientists. Stephan and Levin (2001) find evidence of the extra vitality brought by foreign scientists (foreign-born and foreign-educated) to the U.S. in the fields of Science and Engineering (S&E). Grogger and Hanson (2015) study the mobility of foreign-born students in S&E after earning an American Ph.D. degree, claiming positive spillover effects for destination countries. The migration of German-affiliated researchers is addressed in Zhao et al. (2021), who find a net outflow of researchers from Germany. The current research keeps the focus on the academia and aims to integrate the knowledge-based mobility literature about the Italian university system. To the best of my knowledge, published papers on Italian scholars have studied the role of individual quality on selection processes (Checchi & Verzillo, 2014; Checchi, De Fraja, & Verzillo, 2014) or its link with the competition and incentives generated within the Italian scientific sector (Checchi, De Fraja, & Verzillo, 2021). There have also been some case studies connecting mobility and human capital of professors (see Abramo, D'Angelo, and Di Costa (2022) for Italy, Ejermo, Fassio, and Källström (2020) for Sweden, and Aksnes et al. (2013) for Norway). These researches focus on the effects of professors' mobility on their perfomance. This project investigates the same relationship, but the other way round: it considers human capital as a possible driver for researchers' mobility. Within the Italian university system, mainly student mobility has been analysed (Agasisti & Bianco, 2007; Bratti & Verzillo, 2019; Triventi & Trivellato, 2008), and no previous works have investigated the drivers of scholars' location choice in Italy.

1.2 Data sampling

1.2.1 Institutional context

Italy is home to the oldest University in Europe[4] and has a long tradition of literates and scholars such as Giovanni Boccaccio, Leonardo da Vinci and Galileo

[4]University of Bologna, founded as a university in 1088.

Galilei, who belong in the upper tail of the human capital distribution. The Italian academic system has interesting peculiarities, which are worth mentioning before the empirical analysis. Italy's education system was centralized for a long time (Cottini, Ghinetti, & Moriconi, 2019), making it subject to the whims of the governing body. This increased the importance of hierarchy within academia, based on informal relationships between the most important chaired scholars and government ministers (Capano, 2008; Rebora & Turri, 2008). In the 20th century, this centralization of the system was intended to reduce the inequalities in the Italian education system (Cottini, Ghinetti, & Moriconi, 2019; Triventi & Trivellato, 2008) and there were some positive outcomes. Social mobility improved (Barone & Guetto, 2016) and performance among geographical areas converged (Baldissera & Cornali, 2020), but academia remains seniority-based (Rebora & Turri, 2008), not only in Italy but throughout Europe (MacLeod & Urquiola, 2021). In 1946, to improve the functionality of the system, the universities' autonomy principle (art. 33 paragraph 6) was defined in the Italian Constitution. This precept aimed to underline local excellence (Checchi & Verzillo, 2014), giving each university the autonomy to hire eligible professors. However, this Constitutional principle entered into force only at the end of the 90s, due to the lack of technical standards. The actual implementation of the reforms[5] fragmented the Italian academic market, and it retained some elements of the seniority-based apparatus (Bertola & Sestito, 2011; Bini & Chiandotto, 2003; Cottini, Ghinetti, & Moriconi, 2019; Rebora & Turri, 2008). Among the other modifications, it is important to note that Berlinguer's decree (DPR n. 390/1998) shifted recruitment from a national to a local process.[6] Some could argue that the decentralization of the recruitment process may have had a negative impact (i.e., more opportunistic behaviour and nepotism (Perotti, 2008)) on the scientific productivity of selected professors. Nevertheless, Battistin, Checchi, and Verzillo (2014) show no significant changes in the quality and meritocracy of the university system after the decentralization in 1998. In 2010,[7] the selection procedure was modified again and became a two-stage process. Nowadays, a scholar has to pass a national open competitive exam to be eligible, and then must win a local contest to be hired by a university (Checchi & Verzillo, 2014; Rossi, 2016).

In a system where seniority was the main driver of an academic career, quality and individual ability may be irrelevant. However, Checchi, De Fraja, and Verzillo (2021) found evidence of the opposite. They showed how the most productive

[5] Among the others, the most important ones are laws n. 382/1980, n. 168/1989, n. 210/1998 and the DPR n. 390/1998.

[6] For a detailed explanation of the recruitment process in Italy see the following web page of the Ministry of Education, Universities and Research (only in Italian): https://www.miur.gov.it/reclutamento-nelle-universita.

[7] Law n. 240/2010.

scholars are those who responded best to an increase in the level of competition within the university sector, even in the presence of weak incentives.

The literature about mobility in Italian academia is thin and mostly focuses on student mobility (Agasisti & Bianco, 2007; Bratti & Verzillo, 2019); I have not found any literature on drivers of professors' location choices. Insight into scholars' mobility within the country, and a characterization of the forces that attract them to an institution, can inform public policy.

1.2.2 Professors and Universities

This research is based on a new dataset. The data collection started with RePEc's[8] ranking of the "Top 25% Institutions and Economists in Italy". I decided to focus on scholars in the economic field to exploit the high quantity and quality of information provided by RePEc website. It uses the EDIRC database (Economics Departments, Institutes, and Research Centers in the World), which includes universities, public agencies, central banks, independent research centres, and associations (for more details see section 2.3 in Zimmermann (2013)). Each institution gains from every author's affiliation RePEc collates, implying an advantage for more populous entities (section 6 in Seiler and Wohlrabe (2011) and Zimmermann (2013)).

For the present work, only the universities in Table 1.10 (Appendix) will be taken into account. In Table 1.10 (Appendix) there are 16 public universities, one polytechnic (UNIVPM), and four privately founded universities (BOCCONI, CATT, FUB and LUISS). I consider as *privately funded* a university not fully funded by the State: BOCCONI and LUISS may be defined as fully private, while CATT and FUB as hybrid, receiving funds from both public and private institutions. I include all of them in the *privately funded* group given their different funding system with respect to *fully public* universities.

Each institution includes a list of members (registered in the RePEc Author Service) and I include these observations in the dataset.[9] The people registered on the server have different roles inside academia. In this study, I only include professors – full, associate, adjunct and assistant – and research fellows (also postdoctoral).[10] I include a few emeritus professors who are still teaching. Only scholars who are active in teaching are included in the sample: I call this a "*teaching disclaimer*" and it captures emeritus professors and academics taking

[8] Research Papers in Economics (RePEc) is a project collecting bibliographic data about papers in economics and similar fields, aiming to spread and enhance relative researches.

[9] The ranking is updated month by month; hence the names collected (and the status granted to them) can change with respect to the period of data collection, which is approximately December 2020 – September 2021.

[10] After determining the status of each member, I exclude doctoral students from the dataset.

part in visiting programs or national/international collaborations. Hence, a visiting professor is only included in the sample if she explicitly mentions her teaching activity at the host university. Scholars "on leave" were not considered part of the sample, given the absence of the *teaching disclaimer*. This rule excluded research centres like CEPR, IZA, CESifo, given the honorific nature of their appointments. Table 1.11 (Appendix) presents the precise taxonomy for the scholars included in the dataset, with quantities and percentages.

Once a scholar is identified, they are associated with their university. This process required a careful investigation for each academic. The Curriculum Vitae (CV) was the main source, but where it was out of date or incomplete I used LinkedIn[11] and personal web pages (institutional and/or private). I used the most updated affiliation at the moment of consultation.[12] Affiliations to telematic universities were not taken into account and research centres were excluded. For those universities with multiple locations, I counted the main location, assuming that the majority of the scholars teaching in one location are also teaching in the other(s). This can generate some bias when locations are far away from each other as in the case of Catholic University, with four locations, in Milan (main building), Brescia, Piacenza, and Rome. I discuss the robustness check for this in the Appendix (section 1.6).

Some scholars are associated with more than one university, in Italy or abroad. Multiple affiliations comprise 7.06% of the sample, with a maximum of four affiliations. In the past, academics linked with multiple institutions were associated with high-quality scores (de la Croix et al., 2023), whereas nowadays it is more common to encounter multiple-affiliated scholars with low bibliometric indicators. Usually, these academics are younger and have a postdoc position in a university while teaching in another institution. Empirically, each affiliation of the same scholar is treated as if it was chosen by different individuals, leading to their overestimation with respect to unique-affiliated scholars (see section 1.3.4). In the following part of the paper, the former will be called repeated movers (RM), and the latter single movers (SM).

Treating multiple affiliations in this way, the initial sample counts 1440 observations. A cleaning process removed from the sample scholars who are no longer members of the Italian academy, Ph.D. students, non-teaching emeritus and visiting professors, and those who are on leave.[13] The cleaning process reduced the sample to 1077 names. This procedure identified 76 universities, of which 39 are foreign universities and 37 are Italian. From this set, universities

[11] Professional social network: linkedin.com

[12] Consultation period: December 2020 - September 2021.

[13] One of the main difficulties was understanding the meaning of the various roles and titles indicated by each scholar. The final dataset was built to the best of available knowledge, however, minor errors may still be present.

with fewer than 20 scholars have been excluded, given their minor relevance for academics' choice and to have balanced choices. In the Appendix (section 1.6), I also show the main results considering a threshold of 5 scholars per university. The different threshold implies less balanced choices but a more comprehensive dataset (Figure 1.1). The resulting list is the set of choices each professor faces when maximising their location decision. With the threshold at 20 scholars, this choice set has 17 universities, all of which are Italian. The number of scholars in the database decreased to 936 observations, the percentage of multiple affiliations is now only 3.10% and the maximum number of associations decreased to three. From here onwards, this is the subset for analysis.[14] Table 1.1 summarises the differences between the original dataset and the subset obtained after dropping universities with fewer than 20 scholars.

TABLE 1.1: Comparison between datasets.

	Original	**Subset**
Tot. observations	1440	936
N. Universities	76	17
Obs. after cleaning process	1077	936
Obs. with known birthplaces	936 (86.91%)	815 (87.07%)
Obs. with known education	1044 (96.94%)	904 (96.58%)
Obs. with not-known education	33	32
Multiple affiliations	7.06%	3.10%
Max n. affiliations	4	3

1.2.3 Data on locations

In my analysis, I study the distance a scholar is willing to travel to a given university to develop her career. I treat distance as an increasing cost for the individual. The further she is from her point of origin, the greater the distance and the higher the cost (Schwartz, 1976), also in terms of family attachment. I collect the birthplace for each observation, and treat this as an observable proxy of scholars' usual life context. Using other locations, like residential locations, would involve some endogeneity issues, which are avoided by using birthplaces (Barbieri, Rossetti, & Sestito, 2011). Other variables are not observable – where academics' families live is non-observable, as is the location of their partner's employment.

[14]This dataset is available upon request, and for privacy reasons the data will be made anonymous.

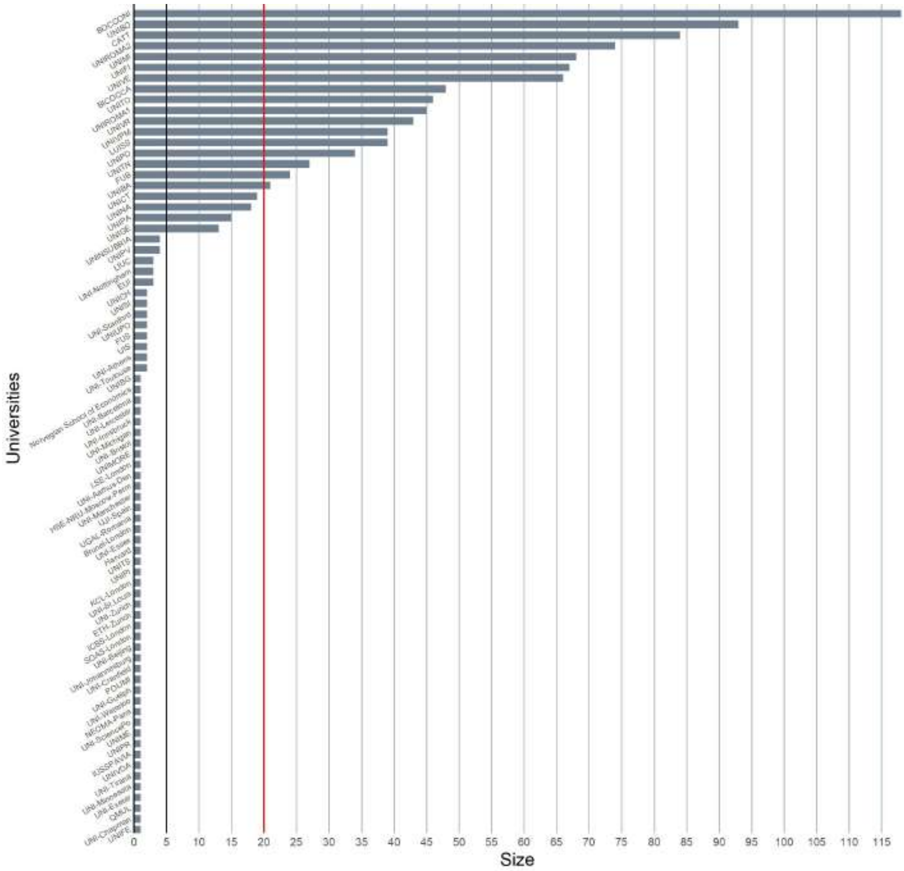


FIGURE 1.1: Histogram of universities' department size. The black line shows the threshold at 5 scholars and the red line shows the threshold at 20 scholars. The universities on the left are those excluded from the analysis at the respective threshold.

CVs and personal webpages were the main sources for affiliations, given that neither LinkedIn nor RePEc provide birthplace information.[15] Only about 30% of the sample indicated their place of birth somewhere in their public profile. Although information about living persons is abundant and often easy to access, bureaucratic and privacy authorizations, which are essential to protect personal information, slow the data collection process. Instead I sent direct emails requesting this information, increasing by about 55% the number of

[15] My thanks to RePEc administrators for their prompt answers.

birthplaces collected.[16] This gives me a known birthplace for 87.07% of the academics (815 observations).

I included in the dataset the location of the institution where each scholar obtained her lowest, publicly-stated degree of education. I consider this another proxy for birthplace, given that the two are likely to coincide or be reasonably close. To deeply investigate this, I study a sub-sample of 789 scholars in my database, for whom I know both their birthplace and their place of education. I consider a scholar to be born and educated in the same cultural environment when the place of birth and education are not further than 60 kilometers.[17] Table 1.2 presents the percentage of scholars who studied, respectively, in the same place where they were born, not further than 30 kilometers and not further than 60 kilometers. More than 60% of this sample was born and studied in the same socio-cultural context. This confirms that the place of education can be used as a proxy for the birthplace.[18] This measure increased the dataset to 904 observations, reaching coverage of 96.58%. For the majority of the sample, the lowest level of education is the bachelor's degree, but some academics mentioned also the high school. Only for a few observations, the lowest education level available was the master's degree, while for five scholars only information about the Ph.D. is known. Given their small number (only 5 out of 936 observations), and given the location of their Ph.Ds: four of them received their doctorate from the same university (or a close one) that they teach at, and only one obtained their title abroad, an ad-hoc robustness check was not necessary. For most observations I found educational information in CVs or LinkedIn profiles, and if I could not find it online I requested it with a direct email.[19] However, education information is missing for 32 observations (3.42% of the sample) and they will be excluded. I implement two different regressions: birthplaces and locations of the lowest level of education analysis (see section 1.3.3).

I match decimal coordinates to location data,[20] giving me a dataset with i observations associated with a geo-localized birth and/or education site, and k geo-localized universities.

1.2.4 Data on quality

For quality indicators I collect *aggregate* quality scores (notability) and *individual* bibliometric indexes (human capital). The former are at the university level,

[16] I express my heartfelt thanks to those who answered in so a interested manner.

[17] 60 kilometers seems a reasonable distance, feasible for commuting every day.

[18] The correlation between distance from birthplaces and distance from educational places is at 60.36%.

[19] Usually in the case of emeritus professors, or persons not in the Italian academia anymore. Thus, this procedure helped also to determine the status of these exceptions.

[20] The websites used are: latitudelongitude.org, tuttitalia.it for Italy, while latlong.net for foreign cities.

TABLE 1.2: Descriptive statistics: scholars with both place of birth and education.

Distance between place of birth and education	Percentage	N. obs out of 789
0 km	41.44	327
30 km	50.32	397
60 km	64.13	506

while the latter are associated with each scholar.

Notability indicators may suffer from endogeneity, because university scores are related to the quality and quantity of scholars. I address this by using past indicators. The average age for Italian academics is 48 years (Morana, 2020) and careers usually begin at around 30 years, so I look for quality indicators from 20 years earlier. The RePEc archives provide aggregate quality scores for top institutions, organised by country, going back as far as 2007. Prior to that, only simple/ordinal rankings are available, and there is no institutional score. I elected to consider scores from around 10 years ago, as of December 2010.[21] Scores for top universities were collected from country rankings. The scores in these rankings are weighted averages of the credit brought by each affiliated scholar: the highest portion (0.5) of affiliation is given to the scholar's main university and the remainder is a weighted average of the other appointments (for the specific formula see section 6 in Zimmermann (2013)). This can generate some biases, for example decreasing the relevance of the main affiliation as more associations are added, as pointed out by Seiler and Wohlrabe (2011).
It was possible to assign a quality score to all 17 universities in the sample. RePEc uses reversed indexes in which lower scores indicate higher quality; I convert them to have a direct relation between indexes and quality. The notability ($\ln Q$) linked to each university ($k \in K$) can be visualized in Figure 1.2 and 1.3.

To compute *human capital* there are many individual bibliometric indicators to choose from. RePEc has the top authors per country ranking (i.e. "Top 25% Institutions and Economists in Italy"). These human capital scores are the harmonic mean of various rankings based on different factors (section 5 in Zimmermann (2013)) and more than 800 scholars are ranked. I use the December 2020 ranking (see below for missing data).[22]

[21]The research started in December 2020.
The ranking is available here: https://ideas.repec.org/top/old/1012/

[22]Given that the ranking is updated every month, the current online score could present some differences.

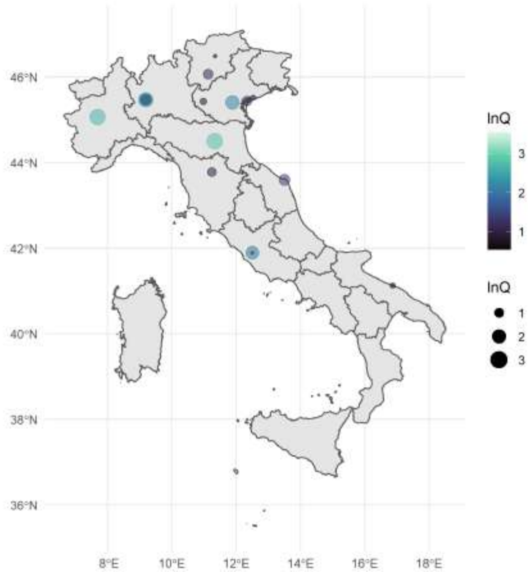


FIGURE 1.2: Bubble Plot on Italy map. Showing notability indexes associated to each university: the higher the $\ln Q$, the bigger and the lighter the colour of the bubble. Each point represents the location of $k \in K$ institutions.
Note: BICOCCA, BOCCONI and CATT are overwritten by UNIMI, which has the highest notability in Milan. UNIROMA1 and LUISS are overwritten by UNIROMA2, which has the highest notability in Rome.

In the literature, academic quality is measured by indicators provided by *Web of Science* (*WoS* – with its three subject specific ISI citation databases; Yang and Meho (2006)). The *WoS* social science indicator goes back to 1956.[23] For a long time it has been one of the few multidisciplinary databases to assign authors' scores based on citations from an original set of sources (Jacso, 2005; Neuhaus & Daniel, 2008). The main issue with Web of Science measures is the relative coverage: only a fraction of sources are considered, although those that are considered (i.e., journal literature) are significant (Norris & Oppenheim, 2007). However, for economics and social science, this literature is not the main way that knowledge is disseminated (Neuhaus & Daniel, 2008).

Quality-evaluation possibilities are now augmented with the automated databases *Scopus*, from Elsevier, and Google Scholar. The former covers a wider

[23]*"Coverage in Web of Science goes back to 1945 for Science Citation Index, 1956 for Social Sciences Citation Index, and 1975 for Arts & Humanities Citation Index."* (Yang & Meho, 2006)

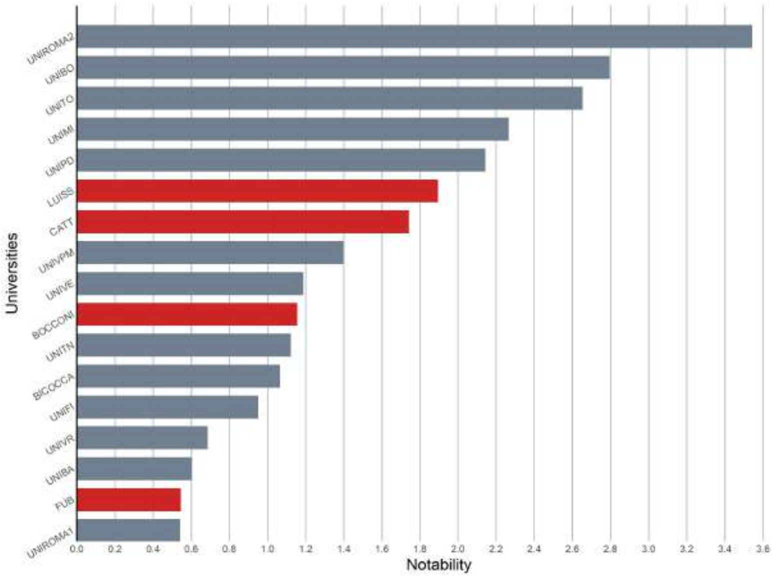


FIGURE 1.3: Histogram of universities' notability. Grey bars define public institutions, red bars private institutions.

range of sources than *WoS*: it starts with an Elsevier database and it goes back to 1996 for social science[24] (Jacso, 2005; Norris & Oppenheim, 2007; Yang & Meho, 2006). *Google Scholar* is a free Google database that uses a wide range of sources, but does not identify clearly what those sources are. This gives it low reliability, which is added to weak, imprecise performance, as pointed out by Neuhaus and Daniel (2008). However, because it is free and has some of the widest coverage among bibliographic indicators, Google scholar still has value as a measure of quality (Neuhaus & Daniel, 2008).

I add to the comparison the *WorldCat identities* index. This database has measures for works (*Worldcat Works*) and library holdings (*Worldcat Library*) for each scholar (and organization) found in WorldCat.org and OCLC sources (OCLC Research, WorldCat identities).[25]

Because no single indicator is perfect, I create a composite indicator of: RePEc score, Worldcat works and library holdings,[26] Google Scholar citations, H-index

[24] Scopus goes back at maximum to 1966. (Yang & Meho, 2006)
[25] https://www.oclc.org/research/areas/data-science/identities.html
[26] https://www.worldcat.org/identities/

and i10-index,[27] WoS H-index,[28] and Scopus H-index.[29] To understand the information added by each indicator I use a *Principal Component Analysis* (PCA) to reduce the number of variables, without losing too much accuracy and information. Once the correlation between the variables is computed (Figure 1.4) and their standardization is completed, the PCA compresses most of the information among the first principal components, which are new uncorrelated variables. For this research, I take the first component into consideration, because its standard deviation is greater than one and the cumulative information explained is sufficiently high (60.78% of the total – Table 1.3). Hence, considering the first component, the analysis gains simplicity while losing only a little portion of its accuracy. The following equation shows the factor loadings. The constant is the minimum of the first component and normalizes it to avoid negative human capital indexes. I use this linear combination of weights to represent the new individual quality index:

$$q_i = + 4.47 - 0.30 \ln(\text{RePEc score}) + 0.32 \ln(\text{Worldcat works}) \\ + 0.30 \ln(\text{Worldcat library holdings}) + 0.37 \ln(\text{Google Scholar citations}) \\ + 0.39 \ln(\text{Google Scholar H-index}) + 0.41 \ln(\text{Google Scholar i10-index}) \\ + 0.35 \ln(\text{WoS H-index}) + 0.37 \ln(\text{Scopus H-index})$$

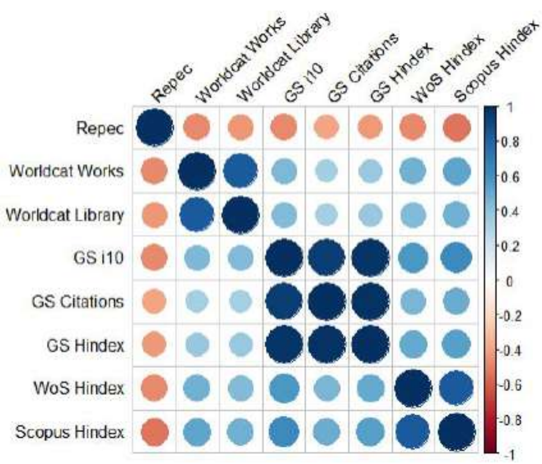


FIGURE 1.4: Correlation Matrix Plot showing the correlations between the eight different bibiliomeric indicators included in the analysis.

[27]https://scholar.google.com/

[28]https://app-webofknowledge-com.pros.lib.unimi.it/author/search?lang=en_US&SID=C45nUucTHDOd2z2VUPs

[29]https://www.scopus.com/freelookup/form/author.uri?zone=TopNavBar&origin=NO%20ORIGIN%20DEFINED

TABLE 1.3: Principal Components table. Showing the standard deviation (St.dv.), the proportion of variance (Pr.Var.) and the cumulative proportion (Cum.Pr.) for each principal component.

	PC1	**PC2**	**PC3**	**PC4**
Sd	2.2051	1.1768	0.9033	0.7604
Var	0.6078	0.1731	0.1020	0.0723
CPr	0.6078	0.7809	0.8829	0.9552
	PC5	**PC6**	**PC7**	**PC8**
Sd	0.4096	0.3861	0.1879	8e−02
Var	0.0210	0.0186	0.0044	8e−04
CPr	0.9762	0.9948	0.9992	1.0000

1.3 Methodology

1.3.1 Main hypotheses

In section 1.2.1, I described some interesting features of Italian universities. One of the main aims of this paper is to understand how these features have changed over time. In order to achieve this objective I compare my results with (de la Croix et al., 2023). The authors tested the following hypotheses for the period between 1000 CE and 1800 CE, while I study them for contemporary times. The role of the location of higher education institutions (Agasisti, Barra, & Zotti, 2019; Audretsch, 1998; Barra & Zotti, 2017; Cottini, Ghinetti, & Moriconi, 2019; Drucker & Goldstein, 2007) is considered to be exogenous.

The current project assumes the presence of strong complementarity in knowledge and skills (Easterly (2001), chapter 8). This property leads to positive assortative matching, where better scholars work together with other high-quality academics. The returns for working with better-skilled personalities are higher for better scholars than for their peers with lower human capital index (Kremer, 1993). Complementarity in knowledge and positive assortative matching may turn into possible virtuous cycles where notable universities, with better scholars, attract more and more high-quality human capital (Easterly, 2001; Kremer, 1993).

Hypothesis 1: *Agglomeration: scholars are attracted by universities with higher notability.*

I expect to find agglomeration (Grogger & Hanson, 2015; Kerr et al., 2016, 2017) although the distance covered by academics could appear shorter than in the past, with a lower magnitude of the coefficient. In Italy, the local appointment of professors may have increased the probability of finding local excellence (Checchi & Verzillo, 2014) – and the importance of networks and nepotism (Durante, Labartino, & Perotti, 2011). With this hypothesis I test for agglomeration forces, such as notability of the university, and the attractiveness of the city in which the institution is located, measured by the size of the population (istat.it) and the local disposable income of private households (finanze.gov.it).

Hypothesis 2: *Positive sorting: scholars with higher human capital weigh the notability of universities higher than scholars with lower human capital do.*

I hypothesise that better scholars have better career prospects, and their expected gains are higher in high-quality environments (Docquier & Marfouk, 2006; Grogger & Hanson, 2015). Thus better professors would assign higher weight to the notability of the university.

Hypothesis 3: *Positive selection: scholars with higher human capital move over greater distances than scholars with lower human capital.*

The literature shows that better-educated people are more mobile (Beine, Bierlaire, & Docquier, 2021; Beine, Docquier, & Özden, 2011; Grogger & Hanson, 2011; Schiller & Cordes, 2016), hence my hypothesis that better professors travel further.

1.3.2 The model

I use a Random Utility Model (RUM), a gravity model widely used in migration analysis (Beine, Bertoli, & Fernández-Huertas Moraga, 2016; Bertoli & Moraga, 2013; Bertoli & Rapoport, 2015; Grogger & Hanson, 2011; Ortega & Peri, 2013). It determines the individual utility of living in a certain region and compares it to the expected utility from moving to alternative locations (Ramos, 2016).

I implement a standard multinomial logit model (Ortega & Peri, 2013), which is a specification of the RUM and requires perfect elasticity of demand in the academic market i.e., that there is a position available for every scholar. In Italian academia, there is a two-step hiring procedure: scholars are filtered at the national level and then at the local level. The assumption of perfectly elastic demand implies that each professor who succeeds at the national level will succeed in finding a chair that she prefers at the local level. This is a reasonable assumption, because the reforms of the university system (in 1998 and in 2010) simplified bureaucratic processes and increased the opening of vacancies (Checchi & Verzillo, 2014; Rossi, 2016). However, in practice only professors with higher individual quality can freely choose the location of their career. To account for this I include the individual human capital score in the analysis. Keeping the perspective of partial equilibrium analyses, I introduce competition variables as demand-side factors, i.e., universities' notability, desirability of the city and individual human capital.

A multinomial logit model allows us to compute the probability that a university k, belonging to the set of choices K, is maximising a scholar i's utility, with error terms independent and identically distributed (D. McFadden, 1974). Technical details are given in the Appendix (section 1.6).

1.3.3 Main results

In this section, I use the multinomial logit model described above to estimate the main regression of the research. First, I consider scholars for whom the place of birth is known (815 observations - 87.07% of the sample). Second, as a robustness check, I use the site of their lowest level of education (904 observations - 96.58% of the sample). I link each site with its geographic coordinates and each academic with a unique individual quality index, computed with a PCA (see section 1.2.4). I discarded universities with fewer than 20 professors from the database, assuming that they have minor relevance in the total set of choices.[30] The university set counts 17 geo-localized institutions linked to their RePEc quality score (see section 1.2.4). Because I work in logarithm terms, the estimation does not allow for zero indexes at aggregate or individual level. If a scholar does not have a positive score, I fill this gap with the lowest human capital index of the sample (794,82 for RePEc, 1 for all the other indicators). It is reasonable to assume that such a scholar does not publish as much as her peers with a positive score. However, it is possible that the sources used to compute bibliometric indicators do not accurately reflect her work, which is a known flaw in quality evaluations. I apply the same reasoning for universities with indexes at zero and link them with the lowest positive score of notability. Finally, I also take the logarithm of the measure of distance which raises the issue of zero distances, affecting scholars born in the same city where they teach. These academics bear the minimum cost of distance, which I assume to be the same as in de la Croix et al. (2023): 3,5 km, the walking distance from the Vatican city to the Colosseum, in the old city of Rome.

In the following part of the section, I describe the results of the main regression which considers scholars' locations of birth. I use the package called "mlogit", written by Croissant (2020). I focus the evaluation on the sign and on the significance of the coefficients of distance, agglomeration, selection, and sorting

[30] In the Appendix I show the main results considering a less stringent threshold of 5 scholars per university - the major difference is in the sorting effect, which falls just below the threshold of significance in the complete model of birthplaces analysis.

effect. I control for unobserved characteristics of universities with fixed effects in each regression, except for when I introduce agglomeration effects. In this case, I include the variables which capture the observed characteristics of the city where the university is located (P_k and Y_k - see section 1.3.2) and represent the reputation of the institution (Q_k). Table 1.4 presents some descriptive statistics.

TABLE 1.4: Descriptive statistics.

Variables	Obs*	Mean	St.Dv.	Min	Max
Birthplace analysis:					
ln of distance	13855	5.454	1.310	1.253	9.198
ln of human capital	13855	4.643	2.157	0.000	12.07
ln of notability	13855	1.544	0.858	0.539	3.541
ln of population	13855	12.99	1.713	7.858	14.88
ln of income	13855	10.16	0.138	9.94	10.39
Lowest level of education analysis:					
ln of distance	15368	5.221	1.569	1.253	9.198
ln of human capital	15368	4.494	2.203	0.000	12.07
ln of notability	15368	1.544	0.858	0.539	3.541
ln of population	15368	12.99	1.713	7.858	14.88
ln of income	15368	10.16	0.138	9.94	10.39

Note: *Obs counts the number of possible dyadic matches in each analysis.

Table 1.5 shows the results of the multinomial logit estimations with known birthplaces. The dataset counts 815 scholars (87.07% of the sample) who choose among 17 universities, resulting in 13855 possible dyadic matches.

The first column contains the basic gravity equation and highlights the negative sign of *distance* coefficients, $\ln d$. This means that the greater the distance between the birthplace and the location of the university, the higher the costs and the lower the probability of finding a dyadic match. Distance coefficients remain highly significant in every specification. The magnitude is consistent with the contemporary migration literature; for example, in *"Diasporas*" (by Beine, Docquier, and Özden (2011)) they also find distance coefficients of around 0.7 when migrants are not divided into low- and high-skilled categories. However, this coefficient is lower than in analyses of past periods (de la Croix et al., 2023).

I add a *selection* effect in the second column, defined by the interaction term between human capital and distance, $\ln q \ln d$. As expected the sign is positive, which means that scholars with higher human capital are less affected by distance than scholars with lower human capital. The high significance of the coefficient (at 1%) confirms the third hypothesis of *positive selection* in every specification of the model.

Column (3) shows the effect of *sorting*, through the interaction between individual human capital and university notability ($\ln q \ln Q$). The positive sign of the coefficient is evidence for *positive sorting*, as expected from the second hypothesis. Despite this, the significance of sorting appears weaker than selection. The sorting effect is non-significant when considered alone in column (3), but it becomes slightly significant (at 10%) in column (4) when I include selection. Sorting maintains the level of significance at 10% in the complete model (column (6)). Finally, I compare log-likelihood (LL) values in order to compute a likelihood-ratio (LR) test: considering column (4) over column (1), the null hypothesis of no selection and no sorting is rejected at any conventional significance level (p-value = 0.000).

To investigate *agglomeration*, I exclude university fixed effects from the regression (columns (5) and (6)), otherwise the effect of agglomeration variables cannot be identified (see section 1.3.2). Without fixed effects, I can study the relevance of the attractiveness of cities where universities are located. All three included variables are highly significant in column (5). The coefficient of the logarithm of population ($\ln P_k$) is negative, which preludes the presence of dispersion: the probability that a scholar chooses university k decreases as the city size increases. The coefficient of the logarithm of disposable income ($\ln Y_k$) is positive, which implies that the variable has a strong attractive force: the richer the city, the greater the likelihood a professor develops her career at that institution. The coefficient of the logarithm of university notability ($\ln Q$) is also significant at 1% and positive, which means that the better the university's reputation, the higher the possibility that a scholar moves there. However, when I consider all the coefficients together (column (6)), notability loses its significance, while the other variables retain their signs, significance levels, and magnitudes – I confirm the second and third hypotheses. The first hypothesis about agglomeration also holds: although from $\ln P_k$ there is a tendency for dispersion (given its negative sign), it is more than compensated by the attractive force of city wealth ($\ln Y_k$). Nevertheless, these results show that agglomeration forces are driven by the income of the city and not by the university's reputation ($\ln Q$). This result reveals room for public policies to improve the relevance of Italian universities' quality in attracting human capital.

1.3.4 Robustness checks

In this section, I substitute the data on scholars' birthplaces with data on their lowest level of education, considered a proxy, which covers 96.58% of the sample.

TABLE 1.5: Multinomial logit regressions: standard logit model - birthplaces analysis, threshold at 20.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.709^{***}	-0.914^{***}	-0.709^{***}	-0.916^{***}	-0.710^{***}	-0.912^{***}
	(0.029)	(0.069)	(0.029)	(0.069)	(0.028)	(0.069)
Selection:						
$\ln q \ln d$		0.043^{***}		0.043^{***}		0.042^{***}
		(0.013)		(0.013)		(0.013)
Sorting:						
$\ln q \ln Q$			0.031	0.032^{*}		0.034^{*}
			(0.019)	(0.019)		(0.020)
Agglomeration:						
$\ln P_k$					-0.122^{***}	-0.127^{***}
					(0.029)	(0.029)
$\ln Y_k$					2.076^{***}	2.179^{***}
					(0.355)	(0.356)
$\ln Q$					0.171^{***}	0.017
					(0.044)	(0.101)
k FE	YES	YES	YES	YES	NO	NO
Obs	815	815	815	815	815	815
R^2	0.157	0.160	0.158	0.161		
LL	−1,867.206	−1,861.567	−1,865.936	−1,860.156	−1,916.167	−1,909.365

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 1.13 (Appendix) presents the results of multinomial logit estimations when I study this proxy. Now the dataset counts 15368 dyadic matches, which associate 904 observations with 17 universities.

Only the *distance* and *agglomeration* coefficients remain significant. The sign of the former is still negative and each specification confirms the magnitude of about 0.7, although it slightly decreases compared to the birthplaces analysis. From the models without fixed effects (columns (5) and (6)), agglomeration variables ($\ln P_k$, $\ln Y_k$, $\ln Q$) confirm again the first hypothesis, with the same signs as in the birthplaces analysis.

The coefficients of *selection effect* ($\ln q \ln d$) are still positive, but not significant anymore. I find similar evidence for *sorting* ($\ln q \ln Q$), which has positive signs but not significant coefficients. These results prove that the second and the third hypotheses are confirmed only when I take into account the actual location of birth; indeed the LR test between (4) and (1) now fails to reject the null hypothesis of no effects (p-value = 0.144). On the other hand, for the standard effect of distance and also for the agglomeration effect, the results remain in line with the birthplace investigation. In this case the analysis focuses on features of the universities (reputation/quality) and cities (population size and income level), aspects that do not vary compared to the previous analysis. The change of dataset affects distance and the individual level of quality, which appear in selection and sorting effects.

Given the results of both regressions, I consider the birthplaces analysis more relevant for the project. I use this as the benchmark model in the following part of the paper, where I develop further analyses.

For space constraints, I elaborate three additional investigations in the Appendix. Firstly, in section 1.6, I correct the human capital index by scholars' age: younger professors with similar bibliometric indicators of senior ones should receive more credit in the computation of their human capital index. Once I introduce age into the analysis, Table 1.16 confirms almost all the benchmark results, but the sorting effect is just below the threshold of significance. These findings indicate that the human capital index employed in the main regression was already able to capture age specificities of Italian scholars. Secondly, I check the overestimation of repeat movers with different strategies in section 1.6. Finally, in section 1.6, I test for gender differences in the effects found in the benchmark model, but no significant discrepancies between male and female professors are found, although women are about one-third of my sample (30.24%).

1.3.5 Private/Public universities

As mentioned in the description of the sample (section 1.2.2), four of the universities originally considered are private: Bocconi University, Catholic University, Free University of Bozen and LUISS University. Private universities have more hiring autonomy and discretion around remuneration (Agasisti & Ricca, 2016; Trivellato, Triventi, & Traini, 2016), making them more attractive to better scholars. To understand how private institutions influence the benchmark estimation, I run additional regressions in the Appendix (one estimation excludes all of them, the others exlude Bocconi and Catholic University one at the time, see section 1.6).

In this section, I develop a nested logit model to investigate further. I divide the set of universities per status s: private and public. The nested logit still denies the correlation of error terms between the two groups (private and public), but there is the possibility of error terms dependency within a nest (D. McFadden, 1978; K. Train, 2003). With this method, it is possible to test whether one type of university implies systematically higher utility. Technical details and general results are found in the Appendix (section 1.6), here I focus on possible differences between private and public institutions.

As the nested logit model is consistent with random utility maximization (details in Appendix), I can define the expected gain each scholar obtains from choosing either a private or a public university. Given the lack of nest-specific variables, this utility is only given by the product $\lambda_s I_{is}$ (explained in Appendix 1.6), which varies for every scholar. Among the 815 professors considered, 127 have greater expected utility (EU) from teaching in public universities than in private ones, while 688 realize higher expected gains by affiliating to private institutions. I compare these two groups (Table 1.6) and the mean of the individual quality for those who prefer public universities is lower (4.02) than for those who prefer private institutions (4.76). *Sorting* effect is evident: better professors prefer more favourable environments. Private institutions, with more available resources, create better contexts to attract more relevant human capital.

TABLE 1.6: Descriptive statistics: groups of scholars preferring public or private universities.

Variables	Obs	Mean	St.Dv.	Min	Max
EU Public > EU Private: ln of human capital	127	4.02	1.86	0.27	7.88
EU Private > EU Public: ln of human capital	688	4.76	2.19	0.00	12.07

1.4 Comparison between the present and the past

It is interesting to compare features of the contemporaneous academic market in Italy with those of the past. I run the same logistic regression as before but I use a sample of professors who worked in Italy from 1000 to 1800,[31] the whole period considered by de la Croix et al. (2023). Agglomeration variables are not fully comparable; I cannot test the first hypothesis with the updated dataset of de la Croix et al. (2023) because the authors consider the level of city democracy instead of the average disposable income of the households ($\ln Y_k$).

In Table 1.7, column (1) summarises the other findings using present professors (i.e., professors who currently work in Italy), while column (2) involves past scholars (i.e., professors who worked in Italy between 1000 and 1800).

In both cases, I confirm standard results for *distance* of gravity models: the greater the distance, the lower the probability a scholar chooses to travel that route. From Table 1.7, the difference in magnitude between these coefficients is evident but both are still in line with the literature, which provides greater magnitude for past periods than for current times. Furthermore, the distance in column (1) is the Euclidean distance, while in de la Croix et al. (2023) it is the cost distance. However, the Euclidean distance increases linearly with the cost distance, which limits the relevance of this computational difference. The magnitude of *selection* effects halves in current times with respect to the past, due to changes in individual quality measures - the human quality indexes are both the result of a *PCA* but they consider different bibliometric indicators.[32] In the Appendix, Table 1.12 compares the total effect of distance now and in the past for different levels of human capital: the effect is almost the same for top scholars in both columns, which confirms the comparability of the results. There is another important difference when I consider *sorting*. To compute the notability of the university, shown in the second column, de la Croix et al. (2023) aggregate the 5 highest human capital indexes associated with scholars active in that institution during the preceding 25 years (for technical details see de la Croix and Stelter (2021)). In the first column, I link the notability index to the RePEc score of each university as of 10 years ago (see section 1.2.4). Finally, the significance of sorting coefficients in Table 1.7 is weaker for current times than for the past, when the same effect had a high relevance.

The time horizon shown in the second column of Table 1.7 is too broad to freely compare it with the shorter time-span of column (1). Instead I exploit the division in periods developed by de la Croix et al. (2023) to seize more directly

[31]Thank you to professor David de la Croix who provided this sample to the project.

[32]Present indicators: RePEc score, Worldcat works and library holdings, Google Scholar citations, H-index and i10-index, WoS H-index, Scopus H-index.
Past indicators: number of characters of the longest Wikipedia page, number of Wikipedia pages in different languages, Worldcat works, library holdings, and publication languages.

Table 1.7: Multinomial logit regressions: standard logit model, birthplaces analysis - comparison of results from the present and from the past without agglomeration variables.

	(1) PRESENT[1]	(2) PAST[2]
Distance:		
$\ln d$	-0.916^{***}	-1.457^{***}
	(0.069)	(0.013)
Selection:		
$\ln q \ln d$	0.043^{***}	0.080^{***}
	(0.013)	(0.005)
Sorting:		
$\ln q \ln Q$	0.032^{*}	0.012^{***}
	(0.019)	(0.002)
k FE	YES	YES
Obs	815	12,280
R^2	0.161	0.399
LL	−1,860.156	−18,627.730

Note: *p<0.1; **p<0.05; ***p<0.01
[1]Present professors, currently working in Italy
[2]Past professors, working in Italy from 1000 to 1800

possible changes and fluctuations of the cycle that occurred in past centuries. In de la Croix et al. (2023) there are eight time segments with a different number of observations available for each period. I follow this two-by-two partition and I group together: the 2nd and 3rd period (1348-1449 / 1450-1526), the 4th and 5th (1527-1617 / 1618-1685), and the 6th and 7th (1686-1733 / 1734-1800). I exclude the first two periods from 1000 to 1347, because there are too few observations, which leads to a negligible empirical relevance and less comparable results with respect to the other segments.

Table 1.8 shows the results. *Distance* is negative and highly significant in every specification. Its magnitude reflects the corresponding literature and historical period: it is higher in columns (2) and (3) than in column (1). This shift corresponds to the rise of national states and the increase in barriers and customs duties, leading to higher transportation costs. These burdens have decreased only in recent times with technological improvements and transport innovations. *Selection* effect is always present, with its positive sign and high relevance. Its magnitude drastically decreases in column (3) and lowers even more when the model involves current scholars.[33] *Sorting* appears the most fluctuating effect across the time horizon, as expected. It is positive and highly significant in the first time range, when most of the major universities are already established (i.e., Bologna, Rome (Sapienza), Florence (Studium generale)) and are among the European top five institutions (de la Croix et al., 2023). These features allow me to position the 1348 - 1526 Italian academic market in the upward part of the aforementioned fluctuating cycle. However, sorting totally disappears in the second period I consider. Its sign is negative in both columns (2) and (3), but it is not significant in either column. These results might be due to the characteristics of the Italian academic world: its decline starts after the sixteenth century, a time of strict censorship of revolutionary concepts by the Catholic Church (Blasutto & de la Croix, 2021). Notable scholars were strongly attracted to the high quality of the first universities, but the sorting effect was diluted with the flow of time and with other universities entering the academic market. This decline in the sorting effect locates the 1527 - 1800 Italian university system in the downward portion of the cycle. Sorting regains its positive sign only when the model considers current scholars. In column (4), positive sorting is slightly significant, which may signal a new momentum for current Italian universities. With the local recruitment of professors and the greater autonomy of each university, quality should gain attention and importance. However, the current analysis cannot detect these reforms with confidence; they are too recent and the influence of the previous seniority-based apparatus persists. This explains the weak sorting effect in the sample of contemporaneous Italian scholars. The same structural explanation applies to the low significance level of sorting in the past (column (2) and (3)): the strong control of the powers in charge (e.g., Catholic Church) limited the relevance of university quality while favouring more denominational sorting, which relies on membership and networks rather than on meritocracy (MacLeod & Urquiola, 2021).

To further emphasize the relevance of *positive sorting* in the functioning of the academic market, I estimate scholars' choice probabilities and compute simulated outcomes with and without the sorting effect. In the Appendix, section 1.6 shows the estimated probabilities for three selected scholars, with different levels of human capital, but born in the same city. Constraining the sorting effect to zero drastically reduces the predicted probability to teach at the best universities. The effect is stronger for better scholars than for professors with lower human capital indexes: for the best scholar (A) the probability to choose the best university (UNIROMA2) halves without the sorting effect. This variation in the choice probabilities is visible for the first six institutions and lowers as I move down in the institutions' ranking. These findings (Table 1.25 - Appendix) demonstrate

[33] When comparing past periods with the present, both human capital indexes are the results of a PCA but they involve different bibliometrics indicators.

TABLE 1.8: Multinomial logit regressions: standard logit model, birthplaces analysis - comparison of results from the present and from the past without agglomeration variables.

	(1)	(2)	(3)	(4)
	(1348-1526)	(1527-1685)	(1686-1800)	(PRESENT)
Distance:				
$\ln d$	-1.360^{***}	-1.609^{***}	-1.598^{***}	-0.916^{***}
	(0.022)	(0.022)	(0.032)	(0.069)
Selection:				
$\ln q \ln d$	0.090^{***}	0.097^{***}	0.056^{***}	0.043^{***}
	(0.010)	(0.009)	(0.013)	(0.013)
Sorting:				
$\ln q \ln Q$	0.014^{***}	-0.0002	-0.010	0.032^{*}
	(0.005)	(0.004)	(0.006)	(0.019)
k FE	YES	YES	YES	YES
Obs.	4472	4643	2314	815
R^2	0.346	0.416	0.435	0.161
LL	$-6,040.509$	$-6,923.913$	$-3,473.547$	$-1,860.156$

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

the importance of positive sorting in fostering high-quality university contexts: the effect is much larger when better professors match with better universities.

To clarify the importance of positive sorting in enhancing quality in academia, I also estimated the total academic output with and without the sorting effect. Section 1.6, in the Appendix, presents the production function I use. I assume the elasticity of substitution between professors' skills, ρ, to be finite. This assumption is crucial because it demonstrates complementarity between professors. As ρ falls, the gains from matching better scholars in the best institutions rise, improving the total output. Table 1.9 presents the results for two levels of ρ: the left part assumes low complementarity ($\rho = 3$), while the right part higher complementarity ($\rho = 2.6$). I estimate the total output by using both the benchmark (Table 1.5 column (6)) and the nested models (Table 1.20 column (6), in the Appendix).

TABLE 1.9: Academic market output - Role of sorting.

	Low complementarity ($\rho = 3$)	
	Benchmark(6)	Nested(6)
With sorting:	26.27	26.58
Without sorting:	25.98	26.25
Δ sorting:	0.28	0.33
	High complementarity ($\rho = 2.6$)	
	Benchmark(6)	Nested(6)
With sorting:	42.93	43.62
Without sorting:	42.32	42.91
Δ sorting:	0.60	0.71

Note: All the values are divided by 1000

When I compare the academic output with and without sorting in Table 1.9, the effect is already clear: when I do not constrain sorting to be zero, the gains are always higher, and this variation is greater with the nested logit model. To fully capture the importance of positive sorting, I compute academic output with different level of complementarity between scholars. When complementarity is higher, there is a significant increase in gains, with output that nearly doubles when the elasticity of substitution decreases by 0.4. This further underlines the relevance of positive sorting in creating a high-quality academic market. This also supports the background concept of this research: strong complementarity between scholars and positive assortative matching may lead to virtuous cycles by increasing the attraction of relevant human capital towards the best environments (Easterly, 2001; Kremer, 1993).

1.5 Conclusions

Using a new sample of contemporaneous scholars, this research confirms and discloses important features of the Italian academic market. Gravity highlights a recurrent effect widely explained in migration literature. Agglomeration forces of Italian universities are driven by the average disposable income of the city where the institution is located and not by universities' notability. This shows room for policy improvements: the quality of institutions is a strong factor for attracting relevant human capital and must be better exploited. Selection effect is also remarkably strong in the benchmark model, which implies that contemporaneous professors travel longer distances when they have greater human capital indexes. Sorting is weaker in this specification, but still significant and positive, which means that notability is more valuable for scholars with a higher individual quality index. Although it is less clear than the others, this last effect might direct the position of Italy in the human capital accumulation cycle. The difference in current and past sorting represents an important initial step for Italian academia: implementing reforms may enhance Italian universities' notability. Policies to improve the quality of high-education institutions would stimulate excellence and in turn, would increase the attractiveness of Italian universities. This would trigger a virtuous circle for the whole economy - improving the sorting effect will feed the system with more resources, attract more remarkable scholars and increase the likelihood of innovations and economic enhancements, as demonstrated by the predicted academic market output. The United States, which has the top universities and research centers, has reaped the benefits of these positive spillovers. Since the early 1900s, sorting has been much stronger in America than it has been in Europe, where centralized systems favored equal growth of high-education institutions while simultaneously preventing the most promising ones from completely exploiting their potential (MacLeod & Urquiola, 2021). The recent reforms in the Italian system might be seen as a watershed moment: the positive achievements reached in terms of equality under a centralized system (Baldissera & Cornali, 2020; Barone & Guetto, 2016) can be bolstered by growing investments in excellence.

Future research can relax the assumption that demand in academia is totally elastic. This would necessitate the use of alternative gravity models, which do not impose the same stringent constraints as the multinomial logit. Gravity models that allow the consideration of both sides of the market can achieve a more complex general equilibrium analysis. Finally, the notability measure can be improved when using the conventional multinomial logit model. This could be accomplished by creating an index similar to de la Croix et al. (2023) to mitigate (if not eliminate) endogeneity issues with RePEc indicators.

Statements and Declarations

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, under grant agreement No 883033 Did elite human capital trigger the rise of the West? Insights from a new database of European scholars.
I also declare that there is no conflict of interest.
The dataset used in this paper is available upon request, and for privacy reasons the data will be made anonymous.

1.6 Bibliography

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2022). The effect of academic mobility on research performance: The case of Italy. *Quantitative Science Studies*, *3*(2), 345–362.

Agasisti, T., & Bianco, A. D. (2007). Determinants of College Student Migration in Italy: Empirical Evidence from a Gravity Approach. *Available at SSRN 1063481*.

Agasisti, T., Barra, C., & Zotti, R. (2019). Research, knowledge transfer, and innovation: The effect of Italian universities' efficiency on local economic development 2006 - 2012. *Journal of Regional Science*, *59*(5), 819–849.

Agasisti, T., & Ricca, L. (2016). Comparing the Efficiency of Italian Public and Private Universities (2007–2011): An Empirical Analysis. *Italian Economic Journal*, *2*(1), 57–89.

Aksnes, D. W., Rørstad, K., Piro, F. N., & Sivertsen, G. (2013). Are mobile researchers more productive and cited than non-mobile researchers? A large-scale study of Norwegian scientists. *Research Evaluation*, *22*(4), 215–223.

Alesina, A., Harnoss, J., & Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, *21*(2), 101–138.

Artige, L., Camacho, C., & de La Croix, D. (2004). Wealth Breeds Decline: Reversals of Leadership and Consumption Habits. *Journal of Economic Growth*, *9*(4), 423–449.

Audretsch, B. (1998). Agglomeration and the location of innovative activity. *Oxford review of economic policy*, *14*(2), 18–29.

Bahar, D., & Rapoport, H. (2018). Migration, knowledge diffusion and the comparative advantage of nations. *The Economic Journal*, *128*(612), F273–F305.

Bahar, D., Rapoport, H., & Turati, R. (2019). Does Birthplace Diversity Affect Economic Complexity? Cross-Country Evidence. *CESifo Working Paper Series 7950, CESifo*.

Baldissera, A., & Cornali, F. (2020). Geography of human capital in italy: A comparison between macro-regions. *Modern Italy*, *25*(3), 261–277.

Barbieri, G., Rossetti, C., & Sestito, P. (2011). The determinants of teacher mobility: Evidence using italian teachers' transfer applications [Special Issue: Economic Returns to Education]. *Economics of Education Review*, *30*(6), 1430–1444.

Barone, C., & Guetto, R. (2016). Verso una meritocrazia dell'istruzione? Inerzia e mutamento nei legami tra origini sociali, opportunità di studio e destini lavorativi in Italia (1920-2009). *Polis*, *30*(1), 5–34.

Barra, C., & Zotti, R. (2017). Investigating the human capital development–growth nexus: Does the efficiency of universities matter? *International Regional Science Review*, *40*(6), 638–678.

Barra, C., & Zotti, R. (2018). The contribution of university, private and public sector resources to Italian regional innovation system (in) efficiency. *The Journal of Technology Transfer*, *43*(2), 432–457.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The quarterly journal of economics*, *106*(2), 407–443.

Barro, R. J. (2001). Human capital and growth. *American economic review*, *91*(2), 12–17.

Battistin, E., Checchi, D., & Verzillo, S. (2014). La riforma di decentramento dei concorsi universitari in italia. *RIV Rassegna Italiana di Valutazione*, 50–79.

Beine, M., Bertoli, S., & Fernández-Huertas Moraga, J. (2016). A Practitioner's guide to Gravity Models of International Migration. *The World Economy*, *39*(4), 496–512.

Beine, M., Bierlaire, M., & Docquier, F. (2021). New York, Abu Dhabi, London or Stay at Home? Using a Cross-Nested Logit Model to Identify Complex Substitution Patterns in Migration [Institute of Labor Economics (IZA)]. *IZA Discussion Papers 14090*.

Beine, M., Docquier, F., & Özden, Ç. (2011). Diasporas. *Journal of Development Economics*, *95*(1), 30–41.

Beine, M., Docquier, F., & Rapoport, H. (2001). Brain drain and economic growth: Theory and evidence. *Journal of development economics*, *64*(1), 275–289.

Bertola, G., & Sestito, P. (2011). A comparative perspective on italy's human capital accumulation. *Bank of Italy Economic History Working Paper No. 6*.

Bertoli, S., & Moraga, J. F.-H. (2013). Multilateral resistance to migration. *Journal of development economics*, *102*, 79–100.

Bertoli, S., & Rapoport, H. (2015). Heaven's swing door: Endogenous skills, migration networks, and the effectiveness of quality-selective immigration policies. *The Scandinavian Journal of Economics*, *117*(2), 565–591.

Bini, M., & Chiandotto, B. (2003). La valutazione del sistema universitario italiano alla luce della riforma dei cicli e degli ordinamenti didattici. *Studi e note di economia*, *2*, 29–61.

Blasutto, F., & de la Croix, D. (2021). Catholic Censorship and the Demise of Knowledge Production in Early Modern Italy. *CEPR Discussion Paper No. DP16409*.

Börsch-Supan, A. (1990). On the compatibility of nested logit models with utility maximization. *Journal of Econometrics*, *43*(3), 373–388.

Bratti, M., & Verzillo, S. (2019). The 'gravity'of quality: Research quality and the attractiveness of universities in italy. *Regional Studies*, *53*(10), 1385–1396.

Capano, G. (2008). Looking for Serendipity: The Problematical Reform of Government within Italy's Universities. *Higher education*, *55*(4), 481–504.

Checchi, D., & Verzillo, S. (2014). Selecting University Professors in Italy: Much Ado about Nothing? *Mimeo, Milano Italy*.

Checchi, D., De Fraja, G., & Verzillo, S. (2014). *And the Winners Are... An Axiomatic Approach to Selection from a Set* (IZA Discussion Papers No. 8344). Institute of Labor Economics.

Checchi, D., De Fraja, G., & Verzillo, S. (2021). Incentives and Careers in Academia: Theory and Empirical Analysis. *The Review of Economics and Statistics*, *103*(4), 786–802.

Cohen, D., & Soto, M. (2007). Growth and human capital: Good data, good results. *Journal of economic growth*, *12*(1), 51–76.

Cottini, E., Ghinetti, P., & Moriconi, S. (2019). Higher Education Supply, Neighbourhood Effects and Economic Welfare. *CESifo Working Paper Series 7483, CESifo*.

Croissant, Y. (2020). Estimation of random utility models in r: The mlogit package. *Journal of Statistical Software*, *95*(11), 1–41.

Daly, A., & Zachary, S. (1978). Improved multiple choice models.

de la Croix, D., Docquier, F., Fabre, A., & Stelter, R. (2023). The Academic Market and the Rise of Universities in Medieval and Early Modern Europe (1000–1800). *Journal of the European Economic Association*, jvad061.

de la Croix, D., & Stelter, R. (2021). Scholars and literati at the university of göttingen (1734–1800). *Repertorium eruditorum totius Europae*, *4*, 1–8.

Docquier, F., & Marfouk, A. (2006). International migration by education attainment, 1990 - 2000. *International migration, remittances and the brain drain*, 151–199.

Docquier, F., Turati, R., Valette, J., & Vasilakis, C. (2020). Birthplace diversity and economic growth: evidence from the US states in the Post-World War II period. *Journal of Economic Geography*, *20*(2), 321–354.

Drucker, J., & Goldstein, H. (2007). Assessing the regional economic development impacts of universities: A review of current approaches. *International regional science review*, *30*(1), 20–46.

Durante, R., Labartino, G., & Perotti, R. (2011). *Academic Dynasties: Decentralization and Familism in the Italian Academia* (NBER Working Papers No. 17572). National Bureau of Economic Research, Inc.

Easterly, W. (2001). *The Elusive Quest for Growth - Economists' Adventures and Misadventures in the Tropics*. The MIT Press, Cambridge - Massachusetts, London - England.

Ejermo, O., Fassio, C., & Källström, J. (2020). Does mobility across universities raise scientific productivity? *Oxford Bulletin of Economics and Statistics*, *82*(3), 603–624.

Faggian, A., & McCann, P. (2009). Universities, agglomerations and graduate human capital mobility. *Journal of Economic and Social Geography*, *100*(2), 210–223.

Grogger, J., & Hanson, G. H. (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics*, *95*(1), 42–57.

Grogger, J., & Hanson, G. H. (2015). Attracting talent: Location choices of foreign-born PhDs in the United States. *Journal of Labor Economics*, *33*(S1), S5–S38.

Handler, H. (2018). Economic links between education and migration: An overview. *Flash Paper Series, Policy Crossover Center Vienna-Europe*, (4), 1–17.

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of economic literature*, *46*(3), 607–68.

Heiss, F. (2002). Structural Choice Analysis with Nested Logit Models. *The Stata Journal*, *2*(3), 227–252.

Jacso, P. (2005). As we may search–Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases. *Current science*, *89*(9), 1537–1547.

Kerr, S. P., Kerr, W., Özden, Ç., & Parsons, C. (2016). Global talent flows. *Journal of Economic Perspectives*, *30*(4), 83–106.

Kerr, S. P., Kerr, W., Özden, Ç., & Parsons, C. (2017). High-skilled migration and agglomeration. *Annual Review of Economics*, *9*, 201–234.

Kling, C. L., & Herriges, J. A. (1995). An empirical investigation of the consistency of nested logit models with utility maximization. *American Journal of Agricultural Economics*, *77*(4), 875–884.

Kremer, M. (1993). The O-Ring Theory of Economic Development. *The Quarterly Journal of Economics*, *108*(3), 551–575.

León, G., & Miguel, E. (2013). *Transportation choices and the value of statistical life* (tech. rep. No. 19494). National Bureau of Economic Research.

Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, *22*(1), 3–42.

MacLeod, W. B., & Urquiola, M. (2021). Why does the United States have the best research universities? Incentives, resources, and virtuous circles. *Journal of Economic Perspectives*, *35*(1), 185–206.

McFadden, D. (1978). *Modeling the choice of residential location.* Transportation Research Record.

McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior*. University of California at Berkeley. Chapter 4.

McFadden, D. (1979). Quantitative methods for analysing travel behaviour of individuals: Some recent developments. In *Behavioural travel modelling* (pp. 279–318). David Hensher; P. Stopher. London : Croom Helm.

Morana, M. (2020). Focus “Il personale docente e non docente nel sistema universitario italiano" - a.a 2019/2020. *Elaborazioni su banche dati MIUR, DGCASIS–Ufficio VI Gestione patrimonio informativo e statistica*.

Nelson, R. R., & Phelps, E. S. (1966). Investment in Humans, Technological Diffusion, and Economic Growth. *The American Economic Review*, *56*(1/2), 69–75.

Neuhaus, C., & Daniel, H.-D. (2008). Data Sources for Performing Citation Analysis: An Overview. *Journal of documentation*, *64*, 193–210.

Norris, M., & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences’ literature. *Journal of informetrics*, *1*(2), 161–169.

Ortega, F., & Peri, G. (2013). The effect of income and immigration policies on international migration. *Migration Studies*, *1*(1), 47–74.

Perotti, R. (2008). *L’università truccata.* Einaudi. Torino, Italy.

Ramos, R. (2016). Gravity models: A tool for migration analysis. *IZA World of Labor*, 239–239.

Rebora, G., & Turri, M. (2008). La governance del sistema universitario in italia: 1989-2008. *Liuc Papers Serie Economia Aziendale*, *32*.

Rossi, P. (2016). Stato giuridico, reclutamento ed evoluzione della docenza universitaria (1975-2015). *4*(1).

Schiller, D., & Cordes, A. (2016). Measuring researcher mobility: A comparison of different datasets and methods with an empirical application of micro-data for the united states and germay. *OECD Blue Sky Forum*.

Schwartz, A. (1976). Migration, Age, and Education. *Journal of Political Economy*, *84*(4, Part 1), 701–719.

Seiler, C., & Wohlrabe, K. (2011). Ranking economists and economic institutions using RePEc: some remarks. *Ifo Working Paper No. 96*.

Stephan, P. E., & Levin, S. G. (2001). Exceptional contributions to US science by the foreign-born and foreign-educated. *Population research and Policy review*, *20*(1), 59–79.

Train, K. (2003, January). *Discrete Choice Method With Simulation*. Cambridge University press.

Train, K. (2009). *Discrete Choice Methods with Simulation.* (2nd ed.). Cambridge University Press.

Train, K. E., McFadden, D. L., & Ben-Akiva, M. (1987). The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Choices. *The RAND Journal of Economics*, *18*(1), 109–123.

Trivellato, P., Triventi, M., & Traini, C. (2016). Private higher education in italy. In *A global perspective on private higher education* (pp. 29–51). Elsevier.

Triventi, M., & Trivellato, P. (2008). Le onde lunghe dell’università italiana. Partecipazione e risultati accademici degli studenti nel Novecento. *Polis*, *22*(1), 85–118.

Williams, R. (2018). Panel data 3: Conditional logit/fixed effects logit models. *University of Notre Dame*, 1–5.

Yang, K., & Meho, L. I. (2006). Citation analysis: a comparison of Google Scholar, Scopus, and Web of Science. *Proceedings of the American Society for information science and technology*, *43*(1), 1–15.

Ye, M., Chen, Y., Yang, G., Wang, B., & Hu, Q. (2020). Mixed Logit Models for Travelers’ Mode Shifting Considering Bike-Sharing. *Sustainability*, *12*(5), 2081.

Zanardello, C. (2022). Market forces in Italian Academia today (and yesterday). *LIDAM Discussion Paper - 2022/01*.

Zhao, X., Aref, S., Zagheni, E., & Stecklov, G. (2021). International Migration in Academia and Citation Performance: An Analysis of German-Affiliated Researchers by Gender and Discipline Using Scopus Publications 1996-2020.

Zimmermann, C. (2013). Academic rankings with repec. *Econometrics*, *1*(3), 249–280.

1.A Additional tables

TABLE 1.10: University sample.

University name	Location	Status
University of Bologna (UNIBO)	Bologna	Public
Catholic University of the Sacred Heart (CATT)	Milan	Private
University of Verona (UNIVR)	Verona	Public
University of Catania (UNICT)	Catania	Public
University of Milan (UNIMI)	Milan	Public
University of Rome - Tor Vergata (UNIROMA2)	Rome	Public
University of Florence (UNIFI)	Florence	Public
University of Venice (UNIVE)	Venice	Public
Polytechnic University of the Marches (UNIVPM)	Ancona	Public
Sapienza University of Rome (UNIROMA1)	Rome	Public
University of Turin (UNITO)	Turin	Public
University of Trento (UNITN)	Trento	Public
University of Naples Federico II (UNINA)	Naples	Public
University of Padua (UNIPD)	Padua	Public
Bocconi University (BOCCONI)	Milan	Private
University of Genoa (UNIGE)	Genoa	Public
University of Palermo (UNIPA)	Palermo	Public
Free University of Bozen (FUB)	Bolzano	Private
University of Bari (UNIBA)	Bari	Public
University of Milan-Bicocca (BICOCCA)	Milan	Public
Luiss University in Rome (LUISS)	Rome	Private

TABLE 1.11: Taxonomy of scholars.

Categories	**Quantity**	**Percentage**
Full professors	420	39%
Associate professors	303	28.13%
Assistant professors	104	9.66%
Adjunct professors	51	4.74%
Research fellows	116	10.77%
Post-doctoral fellows	30	2.79%
Emeritus professors	8	0.74%
Visiting professors	45	4.18%
Total	**1077**	**100%**

TABLE 1.12: Effect of distance in the present and in the past without agglomeration variables.

	(1) PRESENT[1]	(2) PAST[2]
$q_{min}{}^{*}$	−0.916	−1.457
$q_{75}{}^{**}$	−0.653	−1.405
$q_{max}{}^{***}$	−0.395	−0.447
Obs	815	12,280

[1]Present professors, currently working in Italy
[2]Past professors, working in Italy from 1000 to 1800
*Minimum q: −4.468 for present professors, 0 for past profs.
**75^{th} quantile of q: 1.611 for present professors, 0.359 for past profs.
***Maximum q: 7.600 for present professors, 12.604 for past profs.

TABLE 1.13: Multinomial logit regressions: standard logit model - lowest level of education analysis - threshold at 20.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.679^{***}	-0.729^{***}	-0.679^{***}	-0.734^{***}	-0.680^{***}	-0.731^{***}
	(0.022)	(0.049)	(0.022)	(0.049)	(0.021)	(0.049)
Selection:						
$\ln q \ln d$		0.011		0.012		0.011
		(0.010)		(0.010)		(0.010)
Sorting:						
$\ln q \ln Q$			0.029	0.030		0.029
			(0.019)	(0.019)		(0.019)
Agglomeration:						
$\ln P_k$					-0.094^{***}	-0.096^{***}
					(0.028)	(0.029)
$\ln Y_k$					0.635^{*}	0.685^{*}
					(0.374)	(0.375)
$\ln Q$					0.137^{***}	0.012
					(0.043)	(0.094)
k FE	YES	YES	YES	YES	NO	NO
Obs	904	904	904	904	904	904
R^2	0.224	0.224	0.225	0.225		
LL	−1,911	−1,910	−1,910	−1,909	−1,947	−1,945

Note: *p<0.1; **p<0.05; ***p<0.01

TABLE 1.14: Multinomial logit regressions: nested logit model - lowest level of education analysis - threshold at 20.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.689^{***}	-0.730^{***}	-0.700^{***}	-0.749^{***}	-0.835^{***}	-0.915^{***}
	(0.114)	(0.125)	(0.114)	(0.127)	(0.084)	(0.103)
Selection:						
$\ln q \ln d$		0.008		0.010		0.015
		(0.009)		(0.010)		(0.011)
Sorting:						
$\ln q \ln Q$			0.031	0.032		0.038
			(0.021)	(0.021)		(0.025)
Agglomeration:						
$\ln P_k$					-0.106^{***}	-0.110^{***}
					(0.038)	(0.039)
$\ln Y_k$					1.073^{**}	1.168^{**}
					(0.491)	(0.504)
$\ln Q$					0.147^{***}	-0.018
					(0.055)	(0.122)
$\lambda_{private}$	1.423^{***}	1.413^{***}	1.457^{***}	1.450^{***}	1.350^{***}	1.332^{***}
	(0.307)	(0.306)	(0.315)	(0.315)	(0.205)	(0.213)
λ_{public}	0.981^{***}	0.985^{***}	0.996^{***}	1.003^{***}	1.247^{***}	1.264^{***}
	(0.168)	(0.169)	(0.169)	(0.171)	(0.124)	(0.128)
k FE	YES	YES	YES	YES	NO	NO
Obs	904	904	904	904	904	904
R^2	0.226	0.226	0.227	0.227		
LL	$-1,907$	$-1,906$	$-1,905$	$-1,905$	$-1,944$	$-1,942$

Note: *p<0.1; **p<0.05; ***p<0.01

1.B Standard Logit Model - technical details

A multinomial logit model allows us to compute the probability that a university k, belonging to the set of choices K, is maximising a scholar i's utility (D. McFadden, 1974). The first step is to define the utility function for each i scholar. It is defined by a deterministic component $V_{ik} = \beta x_{ik}$, capturing average benefits and costs of each location choice, and by a random component ϵ_{ik} orthogonal to βx_{ik}, which describes unobservable factors that may influence the utility. The utility function can be written as follows:

$$U_{ik} = V_{ik} + \epsilon_{ik} = \beta x_{ik} + \epsilon_{ik} \tag{1.1}$$

The standard logit model relies on the assumption of independent individual choices, which requires ϵ_{ik} be independent and identically distributed (Extreme Value distributed - type I). Under this assumption, the main equation of the multinomial logit model defines the probability of choosing a university k, which depends on the specificities of that institution compared to the specificities of the remaining set of available choices:

$$p_{ik} \equiv Prob[U_{ik} = \max_{k' \in K} U_{ik'}] = \frac{\exp(\beta x_{ik})}{\sum_{k' \in K} \exp(\beta x_{ik'})} \tag{1.2}$$

Another important assumption of the logit model is the Independence of Irrelevant Alternatives (IIA). With independent and identically distributed error terms, the IIA assumption implies that the choice between two specific alternatives should depend only on their own features, without any influence from a third feature (D. McFadden, 1974). This means that the choice between two universities depends only on the two institutions considered and not on other alternatives. In Sections 1.3.4 and 1.3.5 in the main text, I relax this assumption.

In the next step, I explicit the deterministic component, which captures the difference between average benefits and average costs of choosing $k \in K$. The benefits are an increasing function of the university's notability Q_k (as defined in the previous Section), and of the attractiveness of the city. Hence, I include the variables P_k and Y_k, representing respectively cities' total population (capturing the size of the city) and households' disposable income[34] (capturing the wealth status) of the city in which university k is located. In addition, I include an interaction term ($q_i Q_k$) to capture the fact that better scholars (with a high individual quality index q_i) gain more from a welcoming environment (i.e., a

[34] Italy (2018, average disposable income IRPEF, city-level): https://www.finanze.gov.it/it/

university with high Q_k). Therefore, the benefits equation is:

$$B_{ik} = a_0 + a_1 Q_k + a_2 P_k + a_3 Y_k + a_4 q_i Q_k \tag{1.3}$$

where $\forall a \in \{a_0, a_1, a_2, a_3, a_4\}$ greater than zero.

The costs are mainly influenced by distance. The greater the distance from the birthplace (or education-place) the higher is the burden of travel, and hence the higher the costs. However, the better an academic (i.e., the higher the q_i), the more she has to gain in certain university environment Q_k, implying a cost reduction by the interaction term $q_i Q_i$. In addition, as shown by the literature (Beine, Bierlaire, & Docquier, 2021; Grogger & Hanson, 2011; Schiller & Cordes, 2016), a better scholar should be willing to move longer distances. Human capital negatively affects distance, so I include the interaction term between distance and individual quality ($d_{ik} q_i$), which may decrease the costs. Therefore, the costs equation is:

$$C_{ik} = b_0 - b_4 q_i Q_i + b_5 d_{ik} - b_6 d_{ik} q_i \tag{1.4}$$

where $\forall b \in \{b_0, b_4, b_5, b_6\}$ greater than zero.

Having defined the equations for the benefits and the costs, I specify the net benefit for each dyadic match (i.e., the association of scholar i with university k) and explicit the deterministic component of the utility function. This is done by subtracting (4) to (3), with the addition of a fixed effect γ_k. The subscript k suggests that fixed effects refer to time-invariant, non-measurable universities' characteristics which may influence their ability to attract human capital.[35] These fixed effects almost perfectly identify the agglomeration variables (Q_k, P_k, Y_k) included in the model, given that both are destination-specific and time-invariant. For this reason, I exclude fixed effects when agglomeration forces are considered (more on this later). The final expression is:

$$\beta x_{ik} \equiv V_{ik} \equiv B_{ik} - C_{ik} = \beta_0 + \beta_1 Q_k + \beta_2 P_k + \beta_3 Y_k + \beta_4 q_i Q_k + \beta_5 d_{ik} + \beta_6 d_{ik} q_i + \gamma_k \tag{1.5}$$

where β is a vector, whose parameters are common to each scholar. Specifically, the constant β_0 in equation (5) is the difference between the two constants in (3) and (4). To define the *agglomeration effect*, I look at $\beta_j \equiv a_j$ with $j = \{1, 2, 3\}$, which represent notability of universities and attractiveness of cities. A positive sign indicates the presence of agglomeration and a negative sign evidences dispersion. I measure the *sorting effect* with $\beta_4 \equiv a_4 + b_4$. A positive coefficient means that the higher the individual quality, the smaller the cost (or the higher the gain) to travel to better universities. The coefficient $\beta_5 \equiv -b_5$

[35]Fixed effects are used to overcome omitted variables biases, they control for unobserved variables which do not change over time. If there is a change over time, they can be inefficient, with large standard errors. (Williams, 2018)

captures the expected effect of distance, considered as a cost, as previously mentioned. $\beta_6 \equiv b_6$ underlines the *selection effect*: when there is a positive sign, better scholars move further.

Equation (5) includes only destination-specific regressors. Human capital (q_i) is always interacted with university-specific variables (i.e., notability and distance), because it influences all dyadic matches in a symmetric manner.

1.C Age analysis

In the benchmark model, I do not consider the age of scholars to compute the human capital index. I can define this as the *current* human capital indicator, computed as today. However, younger professors with the same level of human capital index as senior ones ought to receive more credit. I include professors' age in the dataset, to compute *age-expected* human capital index. CVs are the main source for this data: most of the scholars disclose their year of birth. However, for 28.85% of the dataset this information is missing and for these cases, I look to the final year of their Ph.D. and assume that scholars at the end of their doctorate are 30 years old. This leaves 29 observations for which I do not have any age reference, neither the year of birth nor the last year of their Ph.D. I exclude them from the analysis, bringing the number of dyadic matches to 13362, which corresponds to 786 academics. Figure 1.5 presents the age distribution of scholars where the average age is 50 years and a few months.

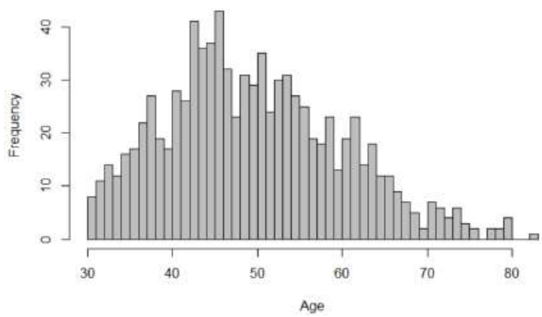


FIGURE 1.5: Histogram showing the age distribution of scholars: mean 50.0751, median 49, variance 108.505 (std. dev. 10.4166), skewness 0.4387 (right-skewed distribution) , kurtosis -0.2510 (platykurtic distribution), min 30, max 83.

I expect age to be a crucial factor in explaining individual human capital: as age increases, the probability of publishing highly-cited researches increases, augmenting in turn the individual quality index. I run an ad-hoc regression, which confirms this expectation at the 1% significance level (Table 1.15). I estimate the *age-expected* human capital index at 40 years old for each scholar using the coefficients of the age regression. I re-run the benchmark model with this new indicator and show the results in Table 1.16.

Distance ($\ln d$), *agglomeration* ($\ln P_k$, $\ln Y_k$, $\ln Q$) and *selection* ($\ln q \ln d$) coefficients have the same sign as in the benchmark model and are still highly significant – agglomeration is driven by city wealth also in the age-adjusted model. *Sorting* ($\ln q \ln Q$) coefficients are positive but not significant, they fall slightly below the threshold of significance. The second hypothesis does not hold anymore: I cannot claim that scholars with a higher human capital index weigh more the quality of universities than those with lower individual quality. The low significance level of both the sorting effect and notability in the agglomeration effect stresses the importance of public policies to improve the quality of the university environment in Italy.

In conclusion, when I introduce age, the model confirms almost all the benchmark results, but the sorting effect is just below the threshold of significance. These findings indicate that the *current* human capital index employed in the main regression was already able to capture age specificities of Italian scholars.

TABLE 1.15: Ordinary Least Square: age regression on human capital.

	$\ln q$
AGE	0.485^{***}
	(0.015)
AGE^2	-0.004^{***}
	(0.0001)
Constant	-13.460^{***}
	(0.382)
Observations	13,632
R^2	0.119
Adjusted R^2	0.119
Residual Std. Error	2.026 (df = 13359)
F Statistic	900.433^{***} (df = 2; 13359)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE 1.16: Multinomial logit regressions: standard logit model, birthplaces age-adjusted analysis - threshold at 20.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.703^{***}	-1.002^{***}	-0.703^{***}	-1.001^{***}	-0.704^{***}	-1.004^{***}
	(0.029)	(0.080)	(0.029)	(0.080)	(0.029)	(0.080)
Selection:						
$\ln q \ln d$		0.058^{***}		0.058^{***}		0.058^{***}
		(0.014)		(0.014)		(0.014)
Sorting:						
$\ln q \ln Q$			0.030	0.031		0.032
			(0.021)	(0.021)		(0.022)
Agglomeration:						
$\ln P_k$					-0.117^{***}	-0.123^{***}
					(0.029)	(0.029)
$\ln Y_k$					2.093^{*}	2.219^{*}
					(0.360)	(0.361)
$\ln Q$					0.174^{***}	0.012
					(0.045)	(0.118)
k FE	YES	YES	YES	YES	NO	NO
Obs	786	786	786	786	786	786
R^2	0.154	0.158	0.155	0.159		
LL	−1,805	−1,797	−1,804	−1,796	−1,853	−1,843

Note: *p<0.1; **p<0.05; ***p<0.01

1.D Multiple Affiliations analyses

After I eliminate universities with fewer than 20 scholars, multiple affiliations count for 3.10% of the sample. Given how they enter the dataset (see section 1.2.2 - main text), the choices of these scholars[36] are overweighted with respect to those of single movers.[37] However, there should not be an over-representation

[36]NB: I call these scholars 'repeated movers', which means that they are associated with more than one university.

[37]NB: by 'single mover' I mean a scholar associated with only one university.

of better scholars against worse ones as in de la Croix et al. (2023), because my dataset associates repeat movers with a low-quality score, with an average human capital of 4.12 against 4.66 of single movers.[38] To confirm that repeat movers do not influence benchmark results, I exclude them from the dataset in columns (3) and (4) of Table 1.18 and then I associate them randomly with one of their affiliations in columns (5) and (6). The main variation with respect to the benchmark model is in the notability coefficient, which becomes negative but remains not significant in both modifications. Significance levels remain almost as in the benchmark model, but in both variations (column (3),(4) and (6)) the sorting effect is now significant at 5% and gains some relevance.

So far, I assumed independent career choices, as required by the IIA assumption in standard logistic models. This is violated when individuals choose more than one alternative at the same time, which is the case when scholars are affiliated to more than one university. I develop a mixed logit model to test for correlated preferences. This version of logistic regression allows to consider the presence of heterogeneous agents. It is similar to the standard model but more flexible: the coefficients are scholar-specific and the utility function includes an additional term which permits correlated choices and the relaxation of the IIA assumption (K. Train, 2009; Ye et al., 2020). Hence, the mixed logit model modifies scholar's utility function as follows:

$$U_{ik} = \beta_i x_{ik} + \eta_i x_{ik} + \epsilon_{ik} \tag{1.6}$$

Where the first term is the general deterministic component, which represents the utility of scholar i who chooses university k. The other two terms capture the unobservable part of the function: η_i is an individual deviation and ϵ_{ik} is a random term as before. I assume these two error terms to be normally distributed.

The mixed logit model, with the η term violating the IIA assumption, requires the integration of the conditional probability by using the joint probability density function, $f(\beta_i|\theta)$; where θ summarises the first and the second moment of the distribution. The vector of β coefficients is assumed to be independent and normally distributed and it is of length N. To obtain the unconditional probability of professor i choosing university k, the following formula applies:

$$\begin{aligned} P_{ik} = E(P_{ik}|\beta_i) = \int_{\beta} (P_{ik}|\beta_i) f(\beta_i|\theta) d\beta = \\ = \int_{\beta_1} \int_{\beta_2} \cdots \int_{\beta_N} (P_{ik}|\beta_i) f(\beta_i|\theta) d\beta_1 d\beta_2 \dots d\beta_N \end{aligned} \tag{1.7}$$

[38] To compute the mean of $\ln q$, I consider 26 observations for repeat movers and 789 observations for single movers.

where

$$(P_{ik}|\beta_i) = \frac{\exp(\beta_i x_{ik})}{\sum_{k' \in K} \exp(\beta_i x_{ik'})} \tag{1.8}$$

is the conditional probability.

Simulations are used to draw the parameters from the β distribution: the unconditional probability is the average of the conditional probabilities computed for each scholar (León & Miguel, 2013; K. Train, 2009; Ye et al., 2020). Table 1.17 presents all the specifications of the mixed logit regression, columns (4) and (6) are those reported in Table 1.18.

With respect to the benchmark model, the magnitude varies for every coefficient. Signs and significance levels of *distance* coefficients confirm the gravity literature and magnitudes increase by almost 0.4. *Agglomeration* is still driven by the city wealth which compensate the negative sign of the population coefficient, and notability does not play any role in attracting scholars, as before. *Selection* effect confirms the third hypothesis with its significance levels at 5%. *Sorting* always appears weaker than the other effects: it is positive, but not significant. The LR test between columns (7) and (1) rejects the benchmark version, but I should notice that there are six additional parameters estimated with the mixed logit (not reported in Table 1.18). Nevertheless, the mixed logit is weaker than the benchmark since it involves simulations and not a maximization. Moreover, the assumption on parameters' distribution is essential to obtain these results, which may change when considering another assumption. The original model remains the benchmark.

TABLE 1.17: Multinomial logit regressions: mixed logit model - birthplaces analysis - threshold at 20.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance						
$\ln d$	-1.141^{***}	-1.509^{***}	-1.139^{***}	-1.501^{***}	-1.075^{***}	-1.343^{***}
	(0.086)	(0.178)	(0.086)	(0.178)	(0.073)	(0.153)
Selection						
$\ln q \ln d$		0.078^{**}		0.077^{**}		0.059^{**}
		(0.031)		(0.031)		(0.028)
Sorting						
$\ln q \ln Q$			0.028	0.028		0.033
			(0.022)	(0.022)		(0.026)
Agglomeration						
$\ln P_k$					-0.152^{***}	-0.159^{***}
					(0.044)	(0.043)
$\ln Y_k$					3.228^{***}	3.305^{***}
					(0.504)	(0.516)
$\ln Q$					0.160^{***}	-0.0004
					(0.054)	(0.129)
k FE	YES	YES	YES	YES	NO	NO
Obs	815	815	815	815	815	815
R^2	0.175	0.177	0.175	0.177		
LL	−1,828	−1,824	−1,827	−1,823	−1,877	−1,871

Note: *p<0.1; **p<0.05; ***p<0.01

Note: The mixed logit involves the six s.d. associated to each coefficient only the s.d. linked to $\ln d$ and $\ln Q$ are signif. different from zero.

TABLE 1.18: Repeat Movers' robustness checks and Mixed logit - birthplace analysis threshold at 20.

	Benchmark		Removing RM		RM linked to 1uni.		Mixed Logit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance:								
$\ln d$	-0.916^{***}	-0.912^{***}	-0.965^{***}	-0.960^{***}	-0.955^{***}	-0.950^{***}	-1.501^{***}	-1.343^{***}
	(0.069)	(0.069)	(0.071)	(0.071)	(0.071)	(0.071)	(0.178)	(0.153)
Agglomeration:								
$\ln P_k$		-0.127^{***}		-0.129^{***}		-0.135^{***}		-0.159^{***}
		(0.029)		(0.029)		(0.029)		(0.043)
$\ln Y_k$		2.179^{***}		2.094^{***}		2.143^{***}		3.305^{***}
		(0.356)		(0.371)		(0.365)		(0.516)
$\ln Q$		0.017		-0.049		-0.013		-0.0004
		(0.101)		(0.105)		(0.103)		(0.129)
Selection:								
$\ln q \ln d$	0.043^{***}	0.042^{***}	0.051^{***}	0.050^{***}	0.050^{***}	0.049^{***}	0.077^{**}	0.059^{**}
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.031)	(0.028)
Sorting:								
$\ln q \ln Q$	0.032^{*}	0.034^{*}	0.042^{**}	0.044^{**}	0.038^{*}	0.041^{**}	0.028	0.033
	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)	(0.020)	(0.022)	(0.026)
k FE	YES	NO	YES	NO	YES	NO	YES	NO
Obs	815	815	763	763	789	789	815	815
R^2	0.161		0.169		0.166		0.177	
LL	$-1,860$	$-1,909$	$-1,724$	$-1,774$	$-1,786$	$-1,838$	$-1,823$	$-1,871$

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

RM = Repeat Movers. Note: The mixed logit involves the six s.d. associated to each coefficient, only the s.d. linked to $\ln d$ and $\ln Q$ is significantly different from zero.

1.E Gender analysis

In the dataset, women are about one-third of the sample (30.24%). In this Section, I test for gender differences in the effects found in the benchmark model.

I estimate the same regression with the addition of an interaction term: each categorical variable interacts with a gender dummy (1 if male, 0 if female). As shown in Table 1.19, none of the interaction terms with the gender dummy ($\ln dxM$, $\ln P_k xM$, $\ln Y_k xM$, $\ln QxM$, $\ln q \ln dxM$, $\ln q \ln QxM$) are significant, revealing no evidence of gender differences. When I include the selection effect in the models (columns (2), (4), and (6)), negative *distance* coefficients show lower magnitudes when only women are considered. When I do not include the selection effect in the model (columns (1), (3), and (5)), distance coefficients have greater magnitudes when I consider only women, with positive men's distance coefficients. Agglomeration coefficients of population and city wealth are lowered if I analyse male scholars, notability is reinforced but it is not significant. Selection and sorting are almost always reinforced when the male portion of the sample is involved, apart in column (2) for selection, and in column (6) for sorting. Most of the coefficients for women are in line with the benchmark model, which confirms that there are no significant gender differences.

TABLE 1.19: Multinomial logit regressions: standard logit model, birthplaces analysis - gender differences threshold at 20.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.709^{***}	-0.916^{***}	-0.709^{***}	-0.917^{***}	-0.710^{***}	-0.911^{***}
	(0.029)	(0.069)	(0.029)	(0.069)	(0.028)	(0.069)
$\ln dxM$	0.036	-0.046	0.037	-0.043	0.036	-0.033
	(0.042)	(0.103)	(0.043)	(0.103)	(0.043)	(0.100)
Selection:						
$\ln q \ln d$		0.043^{***}		0.044^{***}		0.042^{***}
		(0.013)		(0.013)		(0.013)
$\ln q \ln dxM$		-0.016		0.016		0.014
		(0.019)		(0.019)		(0.019)
Sorting:						
$\ln q \ln Q$			0.031	0.032^{*}		0.034^{*}
			(0.019)	(0.019)		(0.020)
$\ln q \ln QxM$			0.002	0.002		-0.017
			(0.010)	(0.010)		(0.023)
Agglomeration:						
$\ln P_k$					-0.123^{***}	-0.127^{***}
					(0.029)	(0.029)
$\ln P_k xM$					0.028	0.026
					(0.031)	(0.031)
$\ln Y_k$					2.082^{***}	2.181^{***}
					(0.355)	(0.356)
$\ln Y_k xM$					-0.568	-0.525
					(0.387)	(0.387)
$\ln Q$					0.171^{***}	0.017
					(0.044)	(0.101)
$\ln QxM$					0.026	0.108
					(0.053)	(0.122)
k FE	YES	YES	YES	YES	NO	NO
Obs.	815	815	815	815	815	815
R^2	0.158	0.160	0.158	0.161		
LL	$-1,866$	$-1,860$	$-1,865$	$-1,859$	$-1,914$	$-1,907$

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
"xM" represents the relative effect when only Male scholars are considered.

1.F Nested Logit Model - Technical Details

The nested logit model enables more appropriate comparison of the two groups of universities and permits the relaxation of the IIA assumption. This method still denies correlation of error terms between the two sectors (private and public), but now there is the possibility of error terms dependency within a nest (D. McFadden, 1978; K. Train, 2003). Hence, the IIA assumption holds within a nest, where the unobserved portions of utility still have the same mean, while the assumption does not hold between nests, where means of the error terms can now differ (Heiss, 2002; K. Train, 2003).

In general, the nested logit model permits grouping of alternatives in nests with similar characteristics, with a certain degree of correlation λ_s. I divide the university set K per status s: private and public. Thus, the utility function of scholar i is decomposed into two parts, plus a random component $\epsilon_{ik'}$. The first portion H_{is} depends only on the nest s, and the other portion $M_{ik'}$ depends on a specific alternative k' within nest s (K. Train, 2003). The new utility function is defined as follows:

$$U_{ik'} = H_{is} + M_{ik'} + \epsilon_{ik'} \tag{1.9}$$

for $k' \in K_s$.

Starting from this decomposed utility function, it is possible to describe the probability of choosing $k \in K_s$ as the product between two probabilities: the conditional probability of choosing k given that the choice of nest K_s has been made (i.e., a standard logit model between alternatives in nest K_s) and the marginal probability of choosing universities in nest K_s (i.e., a standard logit model between nests). The probability of the final choice k for scholar i is the product of two standard logit models:

$$p_{ik} = P_{ik|K_s} P_{iK_s} \tag{1.10}$$

where

$$P_{ik|K_s} = \frac{\exp(M_{ik'}/\lambda_s)}{\sum_{k' \in K_s} \exp(M_{ik'}/\lambda_s)} \tag{1.11}$$

and where

$$P_{iK_s} = \frac{\exp(H_{is} + \lambda_s I_{is})}{\sum_{l=1}^{S} \exp(H_{is} + \lambda_l I_{il})} \tag{1.12}$$

with

$$I_{is} = \ln \sum_{k' \in K_s} \exp(M_{ik'}/\lambda_s) \tag{1.13}$$

The quantity I_{is} is called *inclusive value* or *log-sum term*. It is essential for connecting information in the upper model (marginal probability) with information in the lower model (conditional probability) and it is defined by the logarithm of the lower model denominator - equation (1.11) (K. Train, 2003). λ_s is called *log-sum coefficient* or *dissimilarity parameter* and it reveals informations about the degree of error terms correlation: the higher the λ_s, the higher the independence (or the lower the correlation) of the unobserved portion of utility. The standard multinomial logit model requires λ_s be equal to 1, which implies complete independent error terms (i.e., zero correlation of error terms) (Heiss, 2002; K. Train, 2003). λ_s captures the substitutability of alternatives: if there is more substitution within than between nests, then λ_s is lower than one, while if substitution is greater between rather than within nests, then λ_s is greater that one (K. E. Train, McFadden, & Ben-Akiva, 1987).

Once I_{is} multiplies λ_s, their product $\lambda_s I_{is}$ represents the extra expected utility of scholar i from choosing the best university in nest K_s. This extra expected utility is added to H_{is}, which defines the expected utility of choosing whatever alternative is in the nest. H_{is} depends on nest-specific variables, which are not present in my analysis. Hence, only the product $\lambda_s I_{is}$ tells the difference in the expected utility of choosing a private or a public university: the higher the $\lambda_s I_{is}$, the higher the gain for the scholar (K. Train, 2003). For a nested logit model to be globally consistent with Random Utility Models, the density function must be non-negative; this condition is always met for dissimilarity parameters within the unite interval (Börsch-Supan, 1990; Kling & Herriges, 1995). When λ_s are larger than one (K. E. Train, McFadden, & Ben-Akiva, 1987), the consistency condition may still hold locally, i.e. for some value of the explanatory variables (Börsch-Supan, 1990; Kling & Herriges, 1995).

Table 1.20 presents the results of the simultaneous nested logit model for the birthplace analysis (Table 1.14 in section 1.6 shows the nested logit model for the lowest level of education analysis).[39] All results still hold qualitatively when compared to the benchmark regression. One cannot compare magnitudes between the benchmark and the nested logit model directly, given the presence of the additional parameters λ_s, but it is possible to analyze meaningful ratios. With *selection* effect in column (6) of Table 1.20, when a scholar has a human capital index of 10 (i.e., top scholar), her distance costs decrease by more that 35% with respect to a scholar with a human capital indicator of 4 (i.e., average scholar). Once I compute this percentage using benchmark coefficients, I can claim that there is no relevant difference between the two models in terms of *selection*: the percentage of cost reduction is almost the same (-35.20% for the

[39] A consistent nested logit model can be computed also sequentially, but this latter method is less efficient than the simultaneous approach currently employed (Heiss, 2002; K. Train, 2003).

TABLE 1.20: Multinomial logit regressions: nested logit model - birthplaces analysis - threshold at 20.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.860^{***}	-1.137^{***}	-0.874^{***}	-1.170^{***}	-1.085^{***}	-1.452^{***}
	(0.170)	(0.238)	(0.172)	(0.245)	(0.144)	(0.218)
Selection:						
$\ln q \ln d$		0.053^{***}		0.055^{***}		0.069^{***}
		(0.019)		(0.019)		(0.022)
Sorting:						
$\ln q \ln Q$			0.041	0.046^{*}		0.056^{*}
			(0.026)	(0.028)		(0.033)
Agglomeration:						
$\ln P_k$					-0.167^{***}	-0.180^{***}
					(0.050)	(0.053)
$\ln Y_k$					4.093^{***}	4.413^{***}
					(0.827)	(0.893)
$\ln Q$					0.191^{***}	-0.055
					(0.073)	(0.162)
$\lambda_{private}$	1.821^{***}	1.871^{***}	1.886^{***}	1.968^{***}	1.635^{***}	1.730^{***}
	(0.553)	(0.552)	(0.579)	(0.590)	(0.354)	(0.378)
λ_{public}	1.185^{***}	1.218^{***}	1.204^{***}	1.249^{***}	1.582^{***}	1.637^{***}
	(0.240)	(0.247)	(0.243)	(0.253)	(0.210)	(0.223)
k FE	YES	YES	YES	YES	NO	NO
Obs	815	815	815	815	815	815
R^2	0.159	0.161	0.159	0.162		
LL	−1,864	−1,858	−1,862	−1,856	−1,902	−1,894

Note: *p<0.1; **p<0.05; ***p<0.01

nested model,[40] −33.87% for the standard model[41]). The *sorting* effect presents some difference, with more inequalities among scholars in the nested than in the standard specification. In column (6) of Table 1.20, when a scholar has a

[40] The cost for a scholar with a human capital index of 4 is: $-1.452 + 4 \cdot 0.069 = -1.176$
The cost for a scholar with a human capital index of 10 is: $-1.452 + 10 \cdot 0.069 = -0.762$
The cost reduction for a better scholar is: $(-0.762 + 1.176)/-1.176 = -0.3520 =$ **−35.20%**

[41] The cost for a scholar with a human capital index of 4 is: $-0.912 + 4 \cdot 0.042 = -0.744$
The cost for a scholar with a human capital index of 10 is: $-0.912 + 10 \cdot 0.042 = -0.492$
The cost reduction for a better scholar is: $(-0.492 + 0.744)/-0.744 = -0.3387 =$ **−33.87%**

human capital index of 10, with the sorting effect her gains are 199%[42] higher than the gains of a scholar with an individual quality indicator of 4. On the other hand, in the benchmark model (Table 1.5 – main text), gains for better scholars are 133%[43] higher than for scholars with lower human capital.[44] I also compare two opposite situations to compute the gain percentage variation: the gains of a better scholar (i.e., with a human capital index of 10) who teaches in a better university (i.e., with a notability index of 3) to the gains of a scholar with a lower individual quality (i.e., with a human capital index of 4) who teaches in a worse university (i.e., with a notability index of 1). In the nested logit model (Table 1.20), the gains for the better scholar are 796%[45] higher than those of her peer with a lower human capital indicator. When I compare these two opposite situations in the benchmark model, the gains for the better scholar who teaches in a better university are 600%[46] higher than for a scholar with a lower human capital index who teaches in a university with lower notability. In general terms, the *sorting* effect is stronger in the nested than the standard logit model.

The nesting procedure seems to be justified: the null hypothesis of no nests is rejected through a log-likelihood ratio test (LL = -1909.4, p-value = 0.000), and the correlation within nests is different from zero (Wald test = 21.636, p-value = 0.000), but the null hypothesis of unique nest elasticity cannot be rejected (Wald test = 0.3037, p-value = 0.5816; LL = -1894.8, p-value = 0.507). This raises questions about the applicability of the grouping strategy I employ here, although dividing private and public universities appears reasonable. To clearly define the pertinence of this nested logit model, it is necessary to look at the additional parameters in the last two rows of the output, the λ_s. Firstly, all dissimilarity parameters exceed unity, which poses another question on the global consistency of this nested model with utility maximization. Indeed, Daly

[42] The gain for a scholar with a human capital index of 4 is: $-0.055 + 4{\cdot}0.056 = 0.169$
The gain for a scholar with a human capital index of 10 is: $-0.055 + 10{\cdot}0.056 = 0.505$
The increase of gains for a better scholar is: $(0.505 - 0.169)/0.169 = 1.988 =$ **198.82%**

[43] The gain for a scholar with a human capital index of 4 is: $0.017 + 4{\cdot}0.034 = 0.153$
The gain for a scholar with a human capital index of 10 is: $0.017 + 10{\cdot}0.034 = 0.357$
The gain increase for a better scholar is: $(0.357 - 0.153)/0.153 = 1.333 =$ **133.33%**

[44] Here, better scholars have a human capital index of 10, while scholars with a lower human capital index have an indicator of 4.

[45] The gain for a scholar with a human capital index of 4 in a university with notability of 1 is: $-0.055 + 4{\cdot}0.056 = 0.169$
The gain for a scholar with a human capital index of 10 in a university with notability of 3 is: $-0.055 \cdot 3 + 10{\cdot}3{\cdot}0.056 = 1.515$
The gain increase for a better scholar in a better university is: $(1.515 - 0.169)/0.169 = 7.964 =$ **796.45%**

[46] The gain for a scholar with a human capital index of 4 in a university with notability of 1 is: $0.017 + 4{\cdot}0.034 = 0.153$
The gain for a scholar with a human capital index of 10 in a university with notability of 3 is: $0.017{\cdot}3 + 10{\cdot}3{\cdot}0.034 = 1.071$
The gain increase for a better scholar in a better university is: $(1.071 - 0.153)/0.153 = 6 =$ **600%**

and Zachary (1978) and D. McFadden (1979) show that obtaining λ_s inside the unit interval is essential for the model to be globally consistent. Nevertheless, when this consistency is relaxed to hold only locally, the dissimilarity parameters can exceed one; as showed by Kling and Herriges (1995). Specifically, two conditions must be checked. (i) The non-negativity of the first-order partial derivatives of the choice probabilities is the first necessary condition, and it is described as follows:

$$\lambda_s \leq U_{1s}(v) \equiv \frac{1}{1 - Q_s(v)} \quad s = 1, S \tag{1.14}$$

with v_k defined as the utility delivered by each alternative and $v \equiv (v_1, \ldots, v_K)$, and where $Q_s(v)$ is the upper model as in equation 1.12:[47]

$$Q_s(v) = P_{iK_s} = \frac{\exp(\lambda_s I_{is})}{\sum_{l=1}^{S} \exp(\lambda_l I_{il})} \tag{1.15}$$

For this first necessary condition to hold, Q_s must be sufficiently large. (ii) The second condition questions the non-positivity constraint on the mixed second-order of the choice probabilities, as follows:

$$\lambda_s \leq \mathrm{U}_{2s}(v) \equiv \frac{4}{3\left[1 - Q_s(v)\right] + \sqrt{\left[1 + 7Q_s(v)\right]\left[1 - Q_s(v)\right]}} \tag{1.16}$$

To define these conditions, I compute Q_s from equation 1.14 and compare it with the λ_s, which are already in the output (Table 1.20). Kling and Herriges (1995) present different approaches to precisely test the consistency of nested models, I follow them and Table 1.21 shows my results.

One possible approach confronts $\hat{\lambda_s}$ with both $\hat{U}_{1s}(\bar{v})$ and $\hat{U}_{2s}(\bar{v})$, where $\bar{v}$ denotes the mean of the indirect utility function. Already this approach seems to highlight the consistency of my nested model, by finding that $\hat{\lambda_s} \leq \hat{U}_{1s}(\bar{v})$ and $\hat{\lambda_s} \leq \hat{U}_{2s}(\bar{v})$. Notwithstanding, another approach investigates at which level of precision the estimated coefficients (the λ_s) are able to reject - or not - the local consistency. Hence, I develop a one-tailed test for each condition. For the first-order condition, I test the null hypothesis $H_{1O} : \lambda_s \leq U_{1s}(\bar{v})$ against the alternative $H_{1A} : \lambda_s > U_{1s}(\bar{v})$, and for the second-order condition I compare the null hypothesis $H_{1O} : \lambda_s \leq U_{2s}(\bar{v})$ to the alternative $H_{1A} : \lambda_s > U_{2s}(\bar{v})$. The last column of Table 1.21 reports the t-ratios of each test statistic ($t_{1,2} \equiv [\hat{\lambda_s} - \hat{U}_{1,2s}(\bar{v})]/Std.Error$); negative coefficients immediately imply that the null hypothesis of local consistency cannot be rejected, which is almost always the case. Only the second-order consistency condition for public universities is rejected,

[47] The term H_{is} is not reported, because of the lack of nest-specific variables.

TABLE 1.21: Consistency tests of NLM with RUM.

First-Order Conditions			
Nest	$\hat{\lambda}_s$	$\hat{U}_{1s}(\bar{v})$	**t-ratio**
Private	1.7300	18.4201	−44.1279
Public	1.6368	2.1878	−2.4687
Second-Order Conditions			
Nest	$\hat{\lambda}_s$	$\hat{U}_{2s}(\bar{v})$	**t-ratio**
Private	1.7300	3.8092	−5.49714
Public	1.6368	1.3067	1.4798

as it was the case of the previous approach. Nevertheless, grouping private and public universities seems appropriate and the first-order consistency condition largely approves this nesting procedure. I consider these results consistent enough with utility maximisation models.

1.G Private/Public universities - further insights

In this section, I develop other estimations to further investigates the role played by private universities.

Table 1.22 presents the results of the regressions I run when I exclude all four private universities from the sample. The total of dyadic matches is now 7774, with 598 observations and 13 universities.
Distance coefficients confirm previous findings, each of them is negative, highly significant and with a magnitude a little larger than the benchmark, but still in line with the literature. *Agglomeration* variables in column (5) confirm the first hypothesis by the positive sign and high significance level of the population size ($\ln P_k$) and university notability ($\ln Q$) coefficients. However, I cannot confirm the presence of agglomeration in the last column, where notability loses its significance. This result is driven by the decreasing in magnitude of the city wealth ($\ln Y_k$), which is not significant neither in the fifth nor in the sixth column – likely because the private universities excluded are located in rich cities: two of the four are in Milan, the city with the highest disposable income. *Selection* ($\ln q \ln d$) has a positive sign, its significance level is at 1% and its magnitude is similar to the benchmark model. *Sorting* coefficients are positive and highly significant as well, their magnitude almost double. Table 1.22 confirms *positive selection* and *positive sorting* and brings evidence for a reinforcement of the latter

effect when only public universities are included in the model, but there is no *agglomeration*.

TABLE 1.22: Multinomial logit regressions: standard logit model, birthplaces analysis - private universities excluded - threshold at 20.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.735^{***}	-0.945^{***}	-0.735^{***}	-0.943^{***}	-0.732^{***}	-0.931^{***}
	(0.032)	(0.079)	(0.032)	(0.079)	(0.031)	(0.077)
Selection:						
$\ln q \ln d$		0.045^{***}		0.044^{***}		0.043^{***}
		(0.015)		(0.015)		(0.015)
Sorting:						
$\ln q \ln Q$			0.070^{***}	0.069^{***}		0.072^{***}
			(0.022)	(0.022)		(0.022)
Agglomeration:						
$\ln P_k$					-0.082^{***}	-0.092^{***}
					(0.031)	(0.031)
$\ln Y_k$					0.216	0.352
					(0.463)	(0.464)
$\ln Q$					0.209^{***}	-0.106
					(0.046)	(0.107)
k FE	YES	YES	YES	YES	NO	NO
Obs.	598	598	598	598	598	598
R^2	0.208	0.211	0.211	0.214		
LL	$-1{,}179$	$-1{,}174$	$-1{,}174$	$-1{,}169$	$-1{,}199$	$-1{,}189$
Note:				*p<0.1; **p<0.05; ***p<0.01		

Following, I study two additional logistic regressions: one excludes only Catholic University while keeping Bocconi University, and another excludes Bocconi University while keeping Catholic University. With the former, I investigate the issue of secondary locations to understand whether it alters previous results. With the latter, I examine the position of excellence recently reached by Bocconi in several rankings - it has lately become the best university in Italy for economics and related fields and this may be due to its well-known ability to attract high-ranked personalities.

Table 1.23 presents the results obtained by excluding Catholic University:

almost all coefficients are significant. *Distance* and *agglomeration* remain as in the benchmark (Table 1.5 in the main text) for what concern signs and significance, the magnitude is similar as well, the income coefficient experiences the largest variation: a drop of 0.198 (from 2.179 to 1.981). *Positive selection* maintains its significance level of 1% in each specification, with a slight decrease in magnitude. *Positive sorting* is again weaker than selection but it gains significance in the third column when it is considered alone (with respect to zero significance level of the coefficient in the benchmark). This implies that there is no relevant bias due to the imprecise geographical coordinates associated with this university. Although I assume that every scholar teaches only in Milan, the benchmark model is not significantly influenced.

Excluding Bocconi from the set of choices allows me to find only significant coefficients which improves the solidity of the results (Table 1.24). *Distance* and *agglomeration* variables remain with the same significance as in the benchmark model. The magnitude of the notability coefficient decreases and the coefficient turns negative, while income's magnitude almost halves but still confirms agglomeration. I find *positive selection* as in the previous estimation without Catholic University. In this regression, there is strong evidence for *positive sorting* in each specification. This model reaches a high significance level, which means that private universities have different features not totally captured by the variables included in the benchmark: further investigations are necessary (i.e., applying different models, like the nested logit model - see section 1.3.5 in the main text). Despite this, I can claim that all the expected features of the contemporaneous academic world described in section 1.3.1 of the main text still hold if I exclude Bocconi University.

TABLE 1.23: Multinomial logit regressions: standard logit model, birthplaces analysis - Catholic University excluded.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.705^{***}	-0.891^{***}	-0.705^{***}	-0.893^{***}	-0.708^{***}	-0.890^{***}
	(0.029)	(0.071)	(0.029)	(0.070)	(0.029)	(0.070)
Selection:						
$\ln q \ln d$		0.039^{***}		0.039^{***}		0.038^{***}
		(0.013)		(0.013)		(0.013)
Sorting:						
$\ln q \ln Q$			0.034^{*}	0.035^{*}		0.036^{*}
			(0.020)	(0.019)		(0.020)
Agglomeration:						
$\ln P_k$					-0.114^{***}	-0.119^{***}
					(0.029)	(0.029)
$\ln Y_k$					1.872^{***}	1.981^{***}
					(0.383)	(0.384)
$\ln Q$					0.171^{***}	0.004
					(0.044)	(0.103)
k FE	YES	YES	YES	YES	NO	NO
Obs	741	741	741	741	741	741
R^2	0.169	0.171	0.170	0.172		
LL	−1,635	−1,630	−1,633	−1,629	−1,683	−1,677

Note: *p<0.1; **p<0.05; ***p<0.01

TABLE 1.24: Multinomial logit regressions: standard logit model, birthplaces analysis - Bocconi University excluded, threshold at 20.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.708^{***}	-0.920^{***}	-0.709^{***}	-0.921^{***}	-0.710^{***}	-0.919^{***}
	(0.029)	(0.073)	(0.029)	(0.072)	(0.029)	(0.072)
Selection:						
$\ln q \ln d$		0.045^{***}		0.046^{***}		0.045^{***}
		(0.014)		(0.014)		(0.014)
Sorting:						
$\ln q \ln Q$			0.062^{***}	0.063^{***}		0.068^{***}
			(0.021)	(0.021)		(0.021)
Agglomeration:						
$\ln P_k$					-0.120^{***}	-0.128^{***}
					(0.029)	(0.029)
$\ln Y_k$					1.154^{***}	1.276^{***}
					(0.397)	(0.399)
$\ln Q$					0.230^{***}	-0.070
					(0.046)	(0.105)
k FE	YES	YES	YES	YES	NO	NO
Obs	714	714	714	714	714	714
R^2	0.174	0.176	0.176	0.179		
LL	−1,578	−1,573	−1,574	−1,568	−1,612	−1,602

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

1.H Role of sorting

To underline the relevance of *positive sorting* in the functioning of the academic market, I estimate scholars' choice probabilities and compute simulated outcomes with and without the sorting effect. I select three scholars (A, B, C) with different levels of human capital (low, medium, and high) but born in the same city (Florence). Table 1.25 displays the values of their predicted probabilities for each location in the choice set. The values on the left part of Table 1.25 refer to the complete model of the benchmark (Table 1.5 column (6), in the main text). The right portion of Table 1.25 includes the predicted probabilities when I set the sorting effect to zero in the same model. I list the institutions by notability to highlight the effect of sorting on the best universities.

Constraining the sorting effect to be zero drastically reduces the predicted probability to teach at the best universities. The effect is stronger for better scholars than for professors with lower human capital indexes: for the best scholar A the probability to choose the best university UNIROMA2 halves without the sorting effect. This variation in the choice probabilities is visible for the first six institutions and lowers as I move down in the institutions' ranking. These findings (Table 1.25) demonstrate the importance of positive sorting in fostering high-quality university contexts: the effect is much larger when better professors match with better universities.

To clarify the importance of positive sorting in enhancing quality in academia, I also estimated the total academic output with and without the sorting effect. I interpret the academic output as a result of high-level teaching and high-impact research. Following de la Croix et al. (2023), I first aggregate the individual quality of every scholar i predicted to teach in a university k to compute institutions' output (characterized as a CES production function), and I aggregate them to compute the total output of the model – with and without the sorting effect. I proceed as follows:

$$Y = \sum_{k} \left(\sum_{i} \hat{p}_{ik} q_i^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \tag{1.17}$$

where $\hat{p}_{ik}$ is the predicted probability for a scholar i to choose university k, as shown in the three examples above. These values weigh scholar's human capital q_i. The additional parameter ρ denotes the elasticity of substitution between individual quality of professors in producing institutions' output. This parameter is crucial because, by assuming it to be finite, it demonstrates complementarity between professors. As ρ falls, the gains from matching better scholars in the best institutions rise, improving the total output. The results are in the main text, Table 1.9.

1.I Main regressions with the threshold at 5 scholars

I develop the two main regressions of the study with varied thresholds in order to further investigate the peculiarities of the current Italian academic market. Hence, rather than limiting myself to universities with more than 20 scholars, I now present the results considering institutions with more than 5 scholars. With this new threshold, I involve in the analysis also smaller universities - Genoa, Catania, Naples, and Palermo. On the one hand, this expands the scope of the research by allowing more universities in the southern part of Italy to participate (i.e., Catania, Naples, and Palermo). On the other hand, these small universities reduce the balance of the choice set, because they appear within the career possibilities of each scholar but they are rarely chosen.

Table 1.26 and Table 1.27 show the results of the birthplaces analysis and the lower level of education analysis, respectively. The latter is similar to the estimation in the main text, with *distance* and *agglomeration* maintaining their high significance levels and with weak evidence for selection effect and no sorting effect. I find the main difference of the new threshold in the birthplace analysis. In Table 1.26 *distance*, *selection* and *agglomeration* correspond to those of the project, while *sorting* lose almost all the significance. It is slightly significant in column (3) when it is considered alone, but it is not significant when the model includes selection and/or agglomeration. This result is in favour of a more balanced choice set: the results gain clearness and precision when the threshold increased to 20 scholars. Keeping smaller universities in the dataset detracts attention from the real scholars' choices, although the inclusion of Southern universities would improve the reach of the project.

TABLE 1.25: Predicted values of individual location choice probabilities - Role of sorting.

	Benchmark (6)			Benchmark (6) - NO sorting		
	A	**B**	**C**	**A**	**B**	**C**
Birthplace	Florence	Florence	Florence	Florence	Florence	Florence
ln of HC	10.097	4.415	2.487	10.097	4.415	2.487
UNIROMA2	6.5%	3.0%	2.1%	3.3%	2.1%	1.8%
UNIBO	11.1%	7.7%	6.3%	7.2%	6.2%	5.5%
UNITO	4.1%	2.1%	1.5%	2.8%	1.7%	1.4%
UNIMI	7.3%	4.3%	3.3%	5.7%	3.7%	3.0%
UNIPD	6.3%	4.0%	3.2%	5.1%	3.5%	3.0%
LUISS	3.6%	2.3%	1.8%	3.2%	2.1%	1.7%
CATT	6.1%	3.9%	3.1%	5.7%	3.7%	3.0%
UNIVPM	3.8%	2.8%	2.4%	4.0%	2.8%	2.3%
UNIVE	5.6%	4.2%	3.5%	6.3%	4.3%	3.6%
BOCCONI	4.9%	3.6%	3.0%	5.6%	3.7%	3.0%
UNITN	3.4%	2.4%	2.0%	3.9%	2.5%	2.0%
BICOCCA	4.8%	3.5%	2.9%	5.6%	3.6%	3.0%
UNIFI	23.5%	49.0%	58.5%	28.8%	51.9%	60.0%
UNIVR	3.0%	2.6%	2.2%	4.0%	2.8%	2.4%
UNIBA	1.3%	0.8%	0.7%	1.7%	0.9%	0.7%
FUB	2.7%	2.1%	1.8%	3.9%	2.4%	1.9%
UNIROMA1	2.2%	1.8%	1.6%	3.1%	2.0%	1.7%

TABLE 1.26: Multinomial logit regressions: standard logit model - birthplaces analysis, threshold at 5.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.777^{***}	-1.000^{***}	-0.777^{***}	-0.994^{***}	-0.774^{***}	-0.984^{***}
	(0.026)	(0.063)	(0.026)	(0.063)	(0.025)	(0.062)
Selection:						
$\ln q \ln d$		0.046^{***}		0.045^{***}		0.044^{***}
		(0.012)		(0.012)		(0.012)
Sorting:						
$\ln q \ln Q$			0.032^{*}	0.028		0.029
			(0.018)	(0.018)		(0.018)
Agglomeration:						
$\ln P_k$					-0.155^{***}	-0.159^{***}
					(0.026)	(0.026)
$\ln Y_k$					2.535^{***}	2.619^{***}
					(0.310)	(0.311)
$\ln Q$					0.208^{***}	0.073
					(0.040)	(0.093)
k FE	YES	YES	YES	YES	NO	NO
Obs	876	876	876	876	876	876
R^2	0.200	0.203	0.200	0.203		
LL	−2,017	−2,009	−2,016	−2,008	−2,070	−2,061

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

TABLE 1.27: Multinomial logit regressions: standard logit model - lowest level of education analysis - threshold at 5.

	(1)	(2)	(3)	(4)	(5)	(6)
Distance:						
$\ln d$	-0.726^{***}	-0.791^{***}	-0.726^{***}	-0.793^{***}	-0.723^{***}	-0.789^{***}
	(0.020)	(0.046)	(0.020)	(0.046)	(0.020)	(0.045)
Selection:						
$\ln q \ln d$		0.014		0.015^{*}		0.015^{*}
		(0.009)		(0.009)		(0.009)
Sorting:						
$\ln q \ln Q$			0.026	0.027		0.026
			(0.018)	(0.017)		(0.017)
Agglomeration:						
$\ln P_k$					-0.127^{***}	-0.129^{***}
					(0.026)	(0.026)
$\ln Y_k$					1.086^{***}	1.135^{***}
					(0.317)	(0.318)
$\ln Q$					0.178^{***}	0.065
					(0.040)	(0.087)
k FE	YES	YES	YES	YES	NO	NO
Obs	969	969	969	969	969	969
R^2	0.259	0.259	0.259	0.260		
LL	−2,069	−2,067	−2,068	−2,066	−2,108	−2,106

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Chapter 2

Early Modern Academies, Universities and Growth*

Knowledge production is central to modern economic growth, but what role did it play in the past? Despite growing interest in the history of human capital, we still know little about how knowledge shaped long-term development in pre-industrial societies. This paper explores the contribution of academies—dynamic, scientifically oriented institutions that emerged across Europe between 1650 and 1800. Drawing on newly assembled data and employing advanced difference-in-differences methods, I show that academies contributed to sustained urban growth. Using individual-level data on scholars, I further demonstrate that while literary academies had limited long-term effects, scientific academies led to persistent gains. I also document positive spillovers: cities near academies experienced faster growth, and the presence of academies improved the quality of existing universities. These findings provide the first empirical evidence of the pivotal role scientific academies played in shaping Europe's long-run economic development.

*You can find the most recent version of the paper here: https://chiarazanardello.github.io/files/zanardello_jmp.pdf. An earlier version of this article was published as LIDAM Discussion Paper IRES 2024/12 accessible at https://sites.uclouvain.be/econ/DP/IRES/2024012.pdf.

2.1 Introduction

Today, we widely recognize the value of knowledge: human capital and innovation drive economic growth and societal progress (Nelson & Phelps, 1966). Skills in science and technology are especially important in modern economies.[2] Yet we still know relatively little about how knowledge and human capital contributed to major technological advances in the past—largely due to limited historical data.[3] While we now have high-quality data to assess the short-term impact of contemporary universities and research institutions (Bianchi & Giorcelli, 2020), the long-term effects of today's investments in human capital remain difficult to measure. One way forward is to look backward. Historical evidence can help us project how current investments may unfold in the future, by learning from patterns observed in the past. In particular, Early Modern Europe saw the creation of new educational institutions—academies and universities—that brought together talented individuals with the aim of producing knowledge that could benefit society. Over the following centuries, Europe emerged as a global scientific leader. In this paper, I introduce new micro-level data and use cutting-edge empirical methods to investigate how academies contributed to this transformation.

I revisit the idea that a relatively small group of highly skilled individuals played a key role in Europe's development before the Industrial Revolution (Mokyr, 2005a; Mokyr & Voth, 2009). Specifically, this paper studies the impact of experimental academies—new types of institutions dedicated to scientific inquiry—and their interaction with more traditional universities. I ask two questions. First, did these academies contribute to the economic development of European cities? Second, how did their presence affect universities? I examine whether universities responded to the presence of academies by modernizing their teaching and organization. I answer these questions by expanding a unique database that tracks scholars active in European educational institutions from 1000 to 1800 (de la Croix, 2021).

Between the 16^{th} and 18^{th} centuries, European intellectual life underwent major change. New scientific ideas and methods emerged. Instead of relying on traditional authorities such as Aristotle or Ptolemy, scholars began to base claims on observation, experimentation, and evidence (Mokyr, 2016). A new kind of institution developed to support this effort: the academy. These academies—distinct from universities founded from the 12^{th} and 13^{th} centuries—provided a more flexible space for experimentation and problem-solving (Applebaum, 2000; McClellan, 1985). They prioritized practical knowledge and

[2] C.f. Barro (1991, 2001), D. Cohen and Soto (2007), and Hanushek and Woessmann (2008)

[3] See Cantoni and Yuchtman (2014), de la Croix et al. (2023), Dittmar and Meisenzahl (2022), and Squicciarini and Voigtländer (2015) for recent empirical studies on this topic.

the improvement of everyday life, often focusing on local needs. This "useful knowledge," as Mokyr describes it, was a key ingredient in Europe's long-run development (Mokyr, 2005a, p.287). I also explore whether the influence of these academies extended beyond their host cities—consistent with the idea that knowledge is a non-rival good that can spread across space (Romer, 1990). By 1800, nearly every major urban center in Europe either hosted an academy or felt the influence of the academy movement (McClellan, 1985). As a robustness check, I present detailed sensitivity analyses to assess whether my results are driven by a few cities or countries which were highly involved with the academy movement.

To study the role of academies across time and space, I use a new and comprehensive dataset on scholars affiliated with both academies and universities, based on the database developed in de la Croix (2021). This dataset includes micro-level information on nearly 80,000 individuals from over 370 institutions across Europe. For each scholar, we document their institutional affiliations, places of birth and death, and field of study. I also enhance the original dataset by improving coverage of academies with a scientific orientation, based on institutions listed in McClellan (1985).[4]

My main outcome variable is population growth at the city level, a standard proxy for historical economic development in a Malthusian context (Ashraf & Galor, 2011; Buringh, 2021). I use difference-in-differences (DID) methods to compare cities that established academies to those that did not. Since academy creation occurred at different times across cities–*staggered* treatment–I rely on recent DID estimators designed to address bias in such designs (Callaway & Sant'Anna, 2021; De Chaisemartin & d'Haultfoeuille, 2024; Sun & Abraham, 2021).

The key identifying assumption is that, in the absence of an academy, cities would have followed similar growth trends. I find that the parallel trends assumption holds during the 1500–1900 period, reducing concerns about reverse causality—that is, that fast-growing cities were more likely to attract academies. However, questions remain about the exogeneity of academy timing. Specifically, unobserved factors might have simultaneously influenced both the decision to create an academy and urban development.

To address this concern, I conducted a detailed historical investigation into the founding of these institutions. I collected extensive qualitative evidence on their origins, including their organizational structures, sources of funding, and the biographies of their founders. This historical record consistently shows that most academies were initiated by groups of scholars motivated by the desire to revitalize scientific inquiry, rather than by rulers or economic elites seeking direct development gains. Political authorities—such as local lords, bishops, or

[4] Access the database at the following link: https://shiny-lidam.sipr.ucl.ac.be/scholars/

monarchs—often supported these initiatives, but typically at a later stage, once the academic community had already laid the groundwork. This timing reduces the likelihood that academies were strategically created in response to economic conditions.

While I cannot claim full exogeneity, I adopt an empirical strategy that progressively addresses endogeneity concerns. I include city fixed effects to control for time-invariant differences across cities, and country-by-time fixed effects to absorb broader national trends. These controls, together with the staggered treatment structure and historical evidence, mitigate—at least partially—concerns about omitted variable bias and reverse causality.

I find that, on average, cities that created an academy experienced 9–15% higher population growth over the next 100 years compared to those that did not. These estimates are based on the period 1500–1900 and are robust across a wide range of specifications.

The type of academy plays a key role in explaining this effect. Leveraging the scholar-level information in the database, I classify academies based on the fields of study of their members. Scientific academies—those where over 50% of members were active in natural sciences or empirical disciplines—had the strongest and most persistent effects, with growth rates 12–15% higher than those of the control group. In contrast, literary academies had weaker and more delayed effects, and their impact was less stable across models.

Academies not only contributed to growth directly—they also appear to have influenced the institutions around them. In cities that already had universities, the arrival of a scientific academy often led to reform. Historical accounts suggest that universities modernized their curricula and teaching methods under the influence of academies (Applebaum, 2000; McClellan, 1985). Using a quality index of universities from de la Croix et al. (2023), I show that university quality increased by 44% on average within 50 years of the creation of a scientific academy. Literary academies did not have the same effect. To my knowledge, this is the first empirical study to measure the influence of academies on universities and their complementarities.

This paper contributes to three literatures. First, in economic history, it improves our understanding of the upper tail of the human capital distribution in Europe before the Industrial Revolution (Mokyr, 2005b, 2016; Mokyr & Voth, 2009; Ó Gráda, 2016). I provide new micro-level data on scholars and institutions, including detailed biographies and institutional histories. I also build on existing work on the long-run effects of educational institutions (Becker & Woessmann, 2009; Cantoni & Yuchtman, 2014; Cinnirella & Streb, 2017; Dittmar & Meisenzahl, 2022; Squicciarini & Voigtländer, 2015). Universities have long been the focus of this literature: from Cantoni and Yuchtman (2014) we know that universities played an important role in mediating uncertainty during Middle Ages, educating judges and public administrators, and de la Croix et al. (2023) determine the strength of universities' quality and professors' skills in moving and locating high-level knowledge across Europe during the Middle Ages, until the eve of the Industrial Revolution. This paper highlights the complementary role of academies in producing useful, applied knowledge (Gage, 1938; Mokyr, 2003). As far as I am aware, the sole study exploring the newly emerging societies at the end of the 18^{th} century is that of Koschnick, Hornung, and Cinnirella (2022), which centers on German economic societies exclusively. I show that academies more broadly—not just economic ones—played a major role in Europe's intellectual and economic transformation. Additionally, while much of the literature focuses on single-country case studies, my research adopts a pan-European perspective– like Benos et al. (2024), Bosker, Buringh, and Van Zanden (2013), and de la Croix et al. (2023) do– offering a broader view of the economic impact of high-level human capital on the eve of the Industrial Revolution (de la Croix et al., 2023; Serafinelli & Tabellini, 2022).

Second, in the economics of innovation, this paper highlights how scientific knowledge mattered even before 1800 (Dittmar, 2019; Koschnick, 2025; Mokyr, 2003, 2005a). Specifically, Hanlon (2022) provides quantitative evidence that the rise of the engineering profession during the first technological breakthrough fundamentally transformed both the process and direction of innovation. My findings also align with Dittmar and Meisenzahl (2022), who shows that German universities contributed to scientific innovation only after implementing more research activities. Similarly, I find that academies, as research-driven institutions, fostered innovation by applying empirical methods and focusing on local problems.

Third, this paper contributes to the economics of education. It is, to my knowledge, the first to study the interaction between different types of higher education institutions with a historical perspective. I show that academies encouraged universities to reform and improve, thereby contributing to city progress through channels beyond direct economic growth.

The rest of this paper is structured as follows. Section 2.2 provides historical context on academies and their relationship with universities. Section 2.3 describes the data. Section 2.4 outlines the empirical methods. Section 2.5 presents the main findings. Section 2.6 examines the effects on university quality. Section 2.7 discusses sensitivity analyses and potential spillover effects. Section 2.8 concludes.

2.2 Historical and institutional context

2.2.1 Universities and academies

Universities were the first wave of higher educational institutions in Medieval and Early Modern Europe.[5] A second major development occurred in the 1650s with the rise of academies and learned societies. These institutions marked a significant shift in the European educational system, serving as a bridge between two traditions: the classical university model and a more modern, practice-oriented approach to scientific learning and dissemination that would gain prominence in the 19th century (McClellan, 1985). Academies represent the roots of what has been described as an *"extraordinary educational breakthrough"* (Schütte, 2007, p.545), a transformation that would continue and evolve throughout the 19th century.[6]

Europe stands out historically for its bottom-up development of higher educational institutions. These institutions, largely absent elsewhere, were predominantly founded on the European continent. In contrast, other regions followed different models. In China, for example, kinship-based institutions such as clans prioritized social stability and security over economic advancement or educational innovation (Chen & Ma, 2022). In the Middle East and North Africa, the madrasa was the dominant form of higher education (Bosker, Buringh, & Van Zanden, 2013); however, it functioned more like a specialized high school than a university in the European sense.

It is important to acknowledge that the scientific method and major innovations were not exclusive to Europe. Significant discoveries originated in other parts of the world, often preceding those made in Europe (Bala, 2006; Needham, 1964). However, this project focuses on Europe as the first region to industrialize and to develop a system of bottom-up higher educational institutions. Figure 2.1 shows the cumulative number of institutions included in the analysis between 1500 and 1800.

The emergence of academies exposed key shortcomings in the institutional structure of universities. In particular, universities were slow to embrace new ideas, especially in the natural sciences. They remained anchored to traditional

[5] Universities are among the earliest institutions of higher learning to employ multiple masters. Prior to their emergence, cathedral and monastic schools dominated the landscape. These typically had only one master and focused on a narrow curriculum, often limited to a single subject (Pixton, 1998). However, some schools—such as the School of Laon—gained significant popularity and attracted students even from abroad, thanks to the depth and reputation of their teaching (Luscombe, 1969).

[6] Although the academies considered in this paper did not serve students directly, their focus on practical research laid the groundwork for technical colleges and universities of applied sciences in various European countries—such as German *Fachschulen* (Pahl & Ranke, 2019) and Danish *tekniske skolers* (Rasmussen, 1969).

curricula, prioritized ancient authorities, and maintained Latin as the primary language of instruction well into the 18th century—traits that signaled cultural conservatism and resistance to innovation.[7]

Academies, by contrast, created dedicated spaces for empirical inquiry and scientific experimentation. Initially, their relationship with universities was tense. For instance, the Society of Haarlem, founded in 1752, only received official recognition in 1761—after the nearby University of Leiden accepted that the Society would not deliver lectures or publish in Latin (Bierens de Haan, 1952). Over time, however, a clearer division emerged: universities remained teaching institutions focused on formal instruction and degree granting, while academies evolved into research institutions where knowledge was actively produced and exchanged (Applebaum, 2000; McClellan, 1985; Pepe, 2008).

The academies' mission centered on the generation and dissemination of "useful knowledge"—science that could improve local living conditions and address practical challenges (Mokyr, 2005a, 2016). Their experimental approach and open engagement with applied problems eventually helped pressure universities into updating their curricula and organizational structures by the end of the Scientific Revolution (Applebaum, 2000).

The case of Turin illustrates these dynamics. While the university of the city, founded in 1404 and later supported by Enlightenment-era rulers, benefitted from several waves of reform, it remained largely focused on classical education with limited engagement in the sciences (Vallauri, 1875; Zanardello, 2022). In 1757, three students of Professor Giovanni Battista Beccaria's—Giuseppe Lagrange, Giuseppe Cigna, and Giuseppe Saluzzo—founded the Scientific Academy of Turin. Their goal was to create a platform for empirical research and applied science, independent from the university and free from its constraints.

The academy, formally recognized in 1783 by King Vittorio Amedeo III, adopted the motto *veritas et utilitas* ("truth and usefulness") and initiated public competitions to solve real-world problems. Its early challenges addressed issues like employment in the silk-textile sector, urban lighting, and agricultural modernization (Accademia delle Scienze di Torino; de la Croix and Zanardello (2021)). These efforts filled a gap left by the university and exemplify how academies carved out a new institutional space for science that emphasized experimentation, application, and local relevance.

Figure 2.8 reflects the lasting impact of this shift, showing a marked increase in university quality in the late 18th century, as measured by scholarly output and the arrival of notable faculty.[8]

[7] Universities did teach some scientific subjects, notably within the *Quadrivium*—arithmetic, geometry, astronomy, and music—but often in a qualitative rather than empirical or applied form (Applebaum, 2000).

[8] For details on how scholarly output and institutional quality are measured, see Section 2.3.

TABLE 2.1: Main differences between universities and academies.

	UNI	**ACAD**
When	From 12^{th}-13^{th} century	From mid-17^{th} century
How	traditional approach mostly *teaching* institutions	experimental approach mostly *research* institutions
Why	learning for its own sake	creating “useful knowledge" (Mokyr, 2005a, p.287)
What	theology, law, logic, and medicine	science, maths, medicine, agriculture, and philosophy
Language	mostly Latin	mostly Vernacular
Finance	Municipality, student fees, Church	Private donations, memberships fees, only a few public funded

Source: Applebaum (2000), McClellan (1985), and Mokyr (2005b)

2.2.2 Academies' characteristics

Having outlined the differences between higher educational institutions, this section describes the main features of academies across European countries.

Academies were closely linked to the intellectual currents of the Scientific Revolution and the rise of “New Science,” with a strong emphasis on experimental methods and empirical data. A key institutional turning point was the founding of the Accademia del Cimento in Florence in 1657—one of the first academies explicitly dedicated to scientific experimentation (McClellan, 1985). Its creation was made possible by local interest and strong patronage: Grand Duke Ferdinand II and his brother Leopold, both disciples of Galileo, sought to promote and apply Galileo's methods in a free and open intellectual environment (Knowles Middleton, 1971; Maylender, 1930).

Inspired by this model, many later academies turned their attention to practical applications of science—in agriculture, industry, commerce, and the betterment of society. Often, these institutions began informally in the private homes of enlightened individuals who hosted small circles of intellectuals. Over time, these gatherings became more formalized, and official recognition (especially through royal patronage) was key to their development, particularly in France, Italy, Germany, and Sweden.

Academies also formed transnational networks through correspondence and the exchange of publications. Scientific journals played a central role in this diffusion of knowledge, offering a more accessible and dynamic alternative to books (McClellan, 1985). Major examples include the *Philosophical Transactions* of the

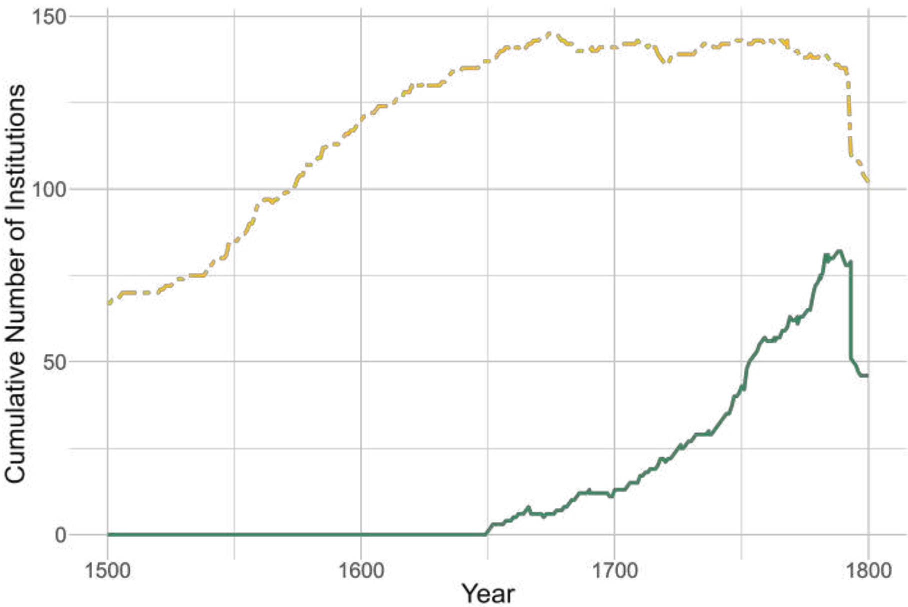


FIGURE 2.1: Cumulative number of institutions over time.

This figure shows the number of universities and academies opened (or closed) over time. The dashed yellow line indicates universities; the solid green line indicates academies. Both openings and closures are included.

Royal Society in London and the *Journal des Savants* which was closely associated with the Académie des Sciences in Paris (henceforth, "Paris Academy"). Other academies published their own periodicals, such as the *Giornale di Letteratura, scienza ed arti* in Messina (now in Italy), the *Monatliche Auszüge* in Olmouc (now in the Czech Republic), and bulletins from academies in Stockholm, Uppsala, and Verona. These publications likely generated spillover effects that extended beyond the immediate location of the academies—analyzed in Section 2.7.3.

Appendix 2.A.1 summarizes key institutional characteristics by modern country, including patterns of official recognition, topics of study, membership, governance, and funding. In France, a highly centralized system ensured that most academies received prompt royal recognition and followed the hierarchical model of the Paris Academy. Italy and Germany showed greater diversity: Italian academies were often influenced by papal support, while German ones leaned more heavily on local patronage. In Great Britain, academies tended to be informal, focused on experimental science, and received limited government funding.

To deepen this institutional overview, I also collected micro-level data on academy founders, who are individuals involved in the initial establishment of the institution. Comparing founders with later affiliates offers insight into the origins of these institutions. Data were compiled for 90 of the 101 academies in my sample. Four of the eleven excluded cases were created directly by monarchs—such as the Academy of Göttingen (King George II of Great Britain) and the Academy of Naples (King Ferdinand IV of Bourbon)—and thus have no individual founder recorded.

Removing those and other cases lacking sufficient data, I identified 413 founders across the sample (as academies often had multiple founders). Table 2.3 compares these individuals with their non-founder peers. Founders appear to have slightly higher individual quality, though this difference is not statistically significant once year fixed effects are included.[9] The two groups do not differ significantly in age at appointment or death, but founders tend to remain active in the academy for longer.

The most striking difference is geographic: founders are more likely to be local. They were born closer to the academy's location and also died closer to their birthplace, indicating lower geographic mobility. This pattern supports the qualitative view that academies were often founded by local scholars motivated to promote science in their own communities. While this paper does not claim full exogeneity in the spatial distribution of academies, these founder characteristics provide suggestive descriptive evidence.

Appendix 2.A.2 lists each academy with a brief account of its founding.

2.3 Data

To assess whether the presence of an academy influenced economic growth in cities, I use data on population, academic institutions, scholars, university quality, and fields of study. This section describes the data sources and construction.

Population data. Population size is a widely used proxy for economic development in pre-industrial societies, based on the assumption that such economies operated under a Malthusian regime; meaning that the higher the technological progress, the larger the population (Ashraf & Galor, 2011). The primary source for population data is Buringh (2021), an updated and extended version of Bairoch, Batou, and Pierre (1988), a widely used historical population dataset. It covers 2,262 cities that reached at least 5,000 inhabitants at some point between 700 and 2000–while Bairoch, Batou, and Pierre (1988) span from 800 to 1850. It also improves on previous efforts by systematically imputing

[9] Individual quality is measured as a composite indicators of citation outputs. See Section 2.3 for more detailed quality measures.

missing values based on city- and time-specific characteristics and correcting earlier miscalculations through new data collection and a refined imputation algorithm. However, some uncertainty remains—particularly in earlier centuries (see Buringh (2021) for details).

Because academies emerged primarily after 1650, I restrict the main analysis to the period after 1500 to minimize potential measurement error in the population data.

The geographical focus is limited to Europe, where the development of educational institutions followed a distinct trajectory. I exclude cities in countries that were part of the Ottoman Empire, removing 161 cities across 11 countries.[10] [11] To avoid inconsistencies and missing values, I aggregate populations of historically separate settlements that later merged into single urban units.[12] The resulting dataset includes population figures for 2,096 cities across 19 time points: every century from 700 to 1400, and every 50 years from 1500 to 2000.

Academies and Universities. While population size is the main dependent variable, my primary independent variables capture the presence of higher education institutions—specifically, academies and universities—reflecting the role of advanced knowledge in a city's development. I rely on a new and unique database on scholars described in de la Croix (2021), which I have expanded to better account for scientific academies. It includes over 79,000 individual scholars active in European universities and academies between 1000 and 1800. For each scholar, the dataset records key information: place and year of birth and death, mobility, affiliations, and a measure of individual quality or "human capital." The database also lists affiliation data: scholars may be affiliated with a university, an academy, or both. For the main analysis, I aggregate these data into binary indicators. The key variable is a dummy for the presence of an academy (ACAD) in city c at time t. Similarly, I define a dummy for the presence of a university (UNI). To capture the effect of institutional overlap, I define a third variable (ACADxUNI), which equals 1 if both an academy and a university are active in the same city and time period, and 0 otherwise. This variable reflects the potential interaction or complementarity between the two types of institutions. The database used in this project covers 79,554 scholars

[10] See Section 2.2.1 for a discussion of the project's regional focus, which acknowledges important scientific developments outside Europe (Bala, 2006; Needham, 1964).

[11] Excluded countries are listed here: https://en.wikipedia.org/wiki/Outline_of_the_Ottoman_Empire (accessed June 2023). Hungary and Slovakia are retained, as they were only partially under Ottoman rule. These exclusions reflect the argument that Ottoman political and strategic priorities may have shaped educational institutions in ways less influenced by the cultural and intellectual currents driving academy formation in Western Europe.

[12] I combine Barmen and Elberfeld into Wuppertal (Germany); Rheydt into Mönchengladbach; Depford into London (Buringh, 2021); and Pest with Buda, which became Budapest in 1873.

and 375 institutions.[13]

For universities, I consider those listed in *A History of the University in Europe* by Frijhoff (1996). I include only institutions identified as "typical" universities, excluding convent-universities, seminaries, collegia, and any that were never operational.[14] This yields data on 171 universities, with location and operating dates based primarily on Frijhoff (1996), supplemented by more precise sources where necessary.[15]

The original database already covers university professors well, so I focus my additions on scholars active in academies.[16] I identify academies based on McClellan (1985), which includes only those with an experimental approach. Renaissance academies are excluded because they followed traditional intellectual models and are not suitable for comparison.[17] I include both official and private academies from McClellan's list, excluding those without corroborating sources.[18] Dates of creation and closure are taken directly from McClellan (1985), unless more accurate sources suggest otherwise. In official cases, I use the date of formal recognition. However, with reliable evidence of activity before that date, I use the foundation year of the society, even if it remained only a private entity for a while.

When a city hosted more than one university or academy, I generally include the oldest institution. This occurs with universities in four cities—Aberdeen, Aix-en-Provence, Nîmes, and Rome. The situation is more complex for academies: several cities had multiple academies (e.g., Bologna, Naples, Florence, and London had three or more). In these cases, I choose the oldest unless another lasted significantly longer, as in Bologna, Caen, and Florence.

Figure 2.2 maps the geographical distribution of educational institutions from 1000 to 1800 CE. Yellow circles represent universities (from Frijhoff (1996)),

---

[13]The version used in this paper is dated January 31, 2025. Institutions with over 100 scholars are documented in the Repertorium Eruditorum Totius Europae: https://ojs.uclouvain.be/index.php/RETE/index.

[14]For example, I exclude the Angelicum in Rome and the University of Camerino (both founded in 1727), which are considered minor or insufficiently documented (Brizzi, 2001).

[15]For the University of Modena, I use information from its official website: https://www.unimore.it/ateneo/cennistorici.html.

[16]This includes both secondary sources and primary material—e.g., handwritten registers from the Academy of Nîmes.

[17]We track only 57 Renaissance academies: 40 in Italy, 12 in France, and one each in Germany, Croatia, the UK, Spain, and Poland. Most (42) were founded before 1650 and few lasted more than a decade—only 25 survived 50 years—making them unsuitable for comparison.

[18]This is the case only in the list of private academies of McClellan (1985, p. 281). I do not consider the Academy *Fisico-matematica* and *dell'Arcidiacono* in Bologna (Italy), the Society in Breman (Germany), the Society in Cuneo (Italy), the Society in Mainz (Germany), the Academy *Clelia de' Vigilanti* in Milan (Italy), the Society in Newcastle-Upon-Tyne (UK), and the *Temple Coffee House Botany Club* of London (UK). I also merged the two academies in Mannheim (Germany): the *Societas Meteorologicae Palatinae* created in 1780 with the *Academia Electoralis Scientiarum et Elegantiorum Literarum Theodoro-Palatina* created in 1763 (Cassidy, 1985).

and green triangles mark academies (from McClellan (1985)). Cities with both institutions—i.e., overlapping markers—represent instances of the ACADxUNI interaction. By 1800, almost every major European city had either an academy or was influenced by the academy movement and its experimental approach (McClellan, 1985).[19] The figure shows only where institutions were founded, not whether they remained open until 1800. Appendix Figure 2.9 illustrates the distribution of founding dates, highlighting the temporal heterogeneity between universities and academies—an important feature for identification. Appendix Figure 2.11 displays the establishment of institutions every 50 years between 1500 and 1800. Finally, I note that measurement error remains a concern: the dataset includes only institutions with sufficient historical visibility, either due to survival or record-keeping.

Scholars. The dataset also includes detailed information on individual scholars, including their VIAF (Virtual International Authority File) identifiers—an online catalogue that tracks name variations, nationalities, publishers, and published works. Additionally, the dataset integrates information from each scholar's Wikipedia page, such as the number of characters (a proxy for entry length) and the number of languages in which the entry is available.

These data are used by de la Croix et al. (2023) to construct a composite measure of scholar quality, referred to as "human capital." This index is derived using principal component analysis (PCA), combining various indicators into a single metric (see Curtis and de la Croix (2023a) for details on the weighting scheme). However, this measure is subject to presentist bias: both VIAF and Wikipedia reflect the current visibility and recognition of historical scholars, rather than their contemporary impact. The metric therefore captures the prominence of those whose work has survived, not necessarily those most influential in their own time.

Language and national biases are also a concern, particularly if relying solely on the English-language Wikipedia, which systematically underrepresents non-Anglophone scholars (Laouenan et al., 2022). To mitigate this, I use Wikipedia entries in all available languages, which significantly reduces this issue. Nonetheless, I acknowledge a broader Western bias, as noted by Laouenan et al. (2022), given the exclusive focus on European institutions. This necessarily excludes significant scholarly contributions from other regions of the world (Bala, 2006; Needham, 1964).

The database further contains individual-level data on scholars' age at death, age at appointment, period of activity in any institution, and geographic mobility—measured by the distance between their birthplace, institutions of activity, and place of death. Table 2.2 reports descriptive statistics for these variables.

[19] Spain has only one academy, in Barcelona; I investigate this further in Section 2.7.1.

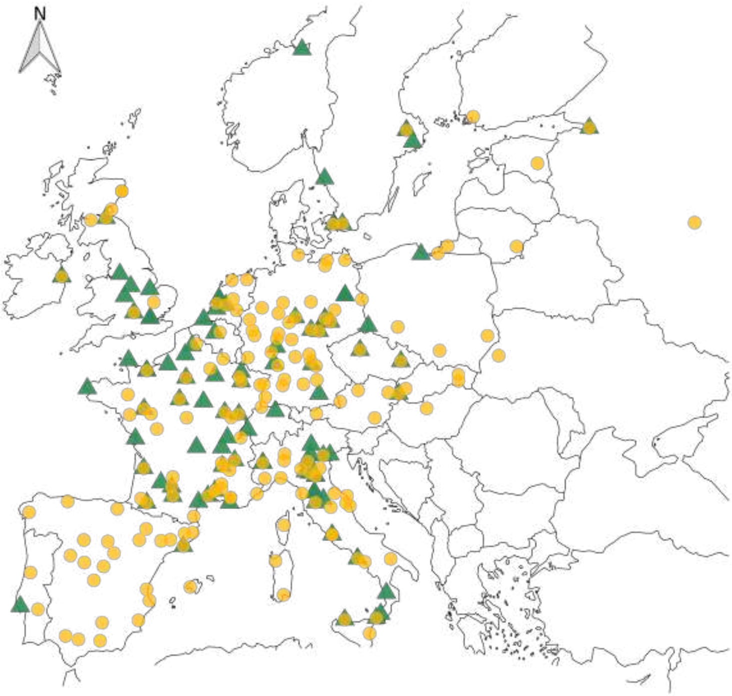


FIGURE 2.2: Locations of higher educational institutions (1000 - 1800 CE).

Yellow circles represent universities, and green triangles represent academies. When both institutions are present in the same city, the overlap of the two shapes indicates an interaction. Country borders reflect those in the year 2000. Only locations of creation are shown; the figure does not reflect whether institutions were still active in 1800.

Column 1 presents data for scholars affiliated with academies, and Column 2 for university professors. To ensure comparability, I restrict the university professor sample to those active from 1600 onwards, excluding earlier medieval scholars. These two groups are distinct, with statistically significant differences across all characteristics (p-values < 0.001).[20] For instance, consistent with findings from Stelter, de la Croix, and Myrskylä (2021), scholars in academies appear to live longer than their university counterparts. Specifically, Stelter, de la Croix, and Myrskylä (2021) document life expectancy trends among members of

[20] I do not report standard errors or coefficients, as all differences are strongly significant.

academies and university professors over time, showing that academicians were the first to experience increased longevity—from the 1750s to the 1850s. They attribute this advantage to social status, arguing that academicians represented an elite subset of scholars with superior academic output and greater access to international networks. This pattern is also reflected in Table 2.2, where academicians display higher individual quality and longer travel distances.

Column 3 shows statistics for individuals affiliated with both a university and an academy at some point in their careers (the "interaction" group). Columns 4 and 5 show the differences between this group and scholars affiliated exclusively with an academy or university, respectively. The interaction group is distinct from both, though it resembles university professors in two aspects: average distance from their institution to their place of death, and from birthplace to place of death. All regressions include year fixed effects to account for secular trends. When scholars are active in multiple institutions simultaneously, I compute the mean value of their individual statistics and count them only once in the sample.

TABLE 2.2: Summary statistics at individual level.

	(1) **ACAD**	(2) **UNI**	(3) **ACADxUNI**	(4) **ACADxUNI** vs ACAD	(5) **ACADxUNI** vs UNI
Number of scholars	16002	32112°	1166		
	μ	μ	μ	p-value	
Quality	2.31	1.01	3.34	0.000	0.000
Age at Death	67	63	68	0.007	0.000
Age at Appointment	38	31	30	0.000	0.000
Activity Years◇	15	10	30	0.000	0.000
Dist. Birth-Institution	338	186	257	0.000	0.000
Dist. Institution-Death	428	188	196	0.000	0.504
Dist. Birth-Death	380	269	247	0.000	0.770
Year FE⋆				YES	YES

Note: Summary statistics are reported for scholars affiliated only with Academies (1), only with Universities (2), and with both Academies and Universities simultaneously at least once in their lifetime (3). Columns (4) and (5) report the statistical significance of the differences between group (3) and groups (1) and (2), respectively. All comparisons control for year fixed effects.
°) University professors are included only if active after 1600.
◇) For activity year statistics, note that when no precise time frame is available in the database: academicians are assumed to be active at their academy for their entire lifetime, while university professors are assumed to be active for 8 years (or until death, if earlier).[21]
⋆) Year fixed effects are based on the scholar's initial year of activity.

University quality. Using the individual-level quality data, I compute an aggregate measure of university quality. Following de la Croix et al. (2023), I adopt their index, which is based on the top five professors who were active at each institution during the 25 years preceding the year for which the quality index is calculated (see de la Croix et al. (2023) for further details). In Section 2.6, I use the 50-year average of this index as the dependent variable.

Fields of study. Universities and academies also differed significantly in their disciplinary focus, as classified in our database (de la Croix & Zanardello, 2022). Figure 2.12 shows the distribution of institutions by main field of study—that is, the discipline with the highest number of affiliated members. Universities primarily concentrated on the humanities, including history, literature, philosophy, ethics, rhetoric, Greek, poetry, theology, and law. Medicine also played a significant role in university curricula. In contrast, academies were predominantly oriented toward the sciences, though 30 academies also had members working in the humanities.

Figure 2.13 displays the distribution of fields by country, separately for academies and universities. Tables 2.4 and 2.5 provide descriptive statistics on size, years of activity, and primary field of study for academies and universities, respectively.

2.4 Empirical strategy

This paper investigates the effects of the creation of early modern academies on urban population growth. The treatment of interest—the founding of an academy—occurred in different cities between 1500 and 1900. To capture its effects, I use a panel dataset with 50-year intervals, focusing on nine such periods that include approximately five intervals before and three after the treatment. Including post-1800 periods is essential to observe long-term effects beyond the initial establishment of academies.

The empirical analysis proceeds in stages, beginning with simpler approaches and progressing toward more robust causal identification strategies. I first employ Ordinary Least Squares (OLS) regressions (Section 2.C.1). While OLS estimates cannot be interpreted causally, they provide preliminary correlational evidence and help assess whether the creation of an academy is systematically associated with changes in population growth. These exploratory results motivate the use of more rigorous methods.

To address endogeneity concerns and uncover dynamic treatment effects, I then use a difference-in-differences (DID) design with a dynamic two-way fixed effects (TWFE) specification (Section 2.C.3). This event-study framework treats academy creation as the focal event and allows for the estimation of both pre- and post-treatment effects. By visualizing treatment effects over time, this method helps identify potential pre-trends and anticipation effects.

For the event-study and subsequent DID estimators to be valid, three key assumptions must hold. First, the *parallel trends assumption* requires that, in the absence of treatment, cities that created an academy would have followed similar growth trajectories to those that did not. Second, the *no anticipation assumption* requires that cities did not change their growth patterns before the academy was officially established. Third, the *Stable Unit Treatment Value Assumption* (SUTVA) implies that the treatment effect in one city is independent of treatment status in neighboring cities. I test the plausibility of these assumptions throughout the analysis. The main results confirm the first two assumptions, and the third is explored further in the robustness checks using spatial analysis and sensitivity exercises (Berkes & Nencka, 2021; Butts, 2021).

Since academies were established at different times across cities, the staggered treatment setting introduces the possibility of treatment effect heterogeneity. Standard TWFE estimators are known to produce biased estimates in such contexts due to inappropriate weighting and implicit assumptions of homogeneity across cohorts (Goodman-Bacon, 2021; Roth et al., 2023). To address this issue, I employ three recent advanced DID estimators that correct for these biases and better account for heterogeneity: the method by Sun and Abraham (2021) (Section 2.5.1), and those developed by De Chaisemartin and d'Haultfoeuille (2024) and Callaway and Sant'Anna (2021) (Appendix 2.C.5). These estimators construct more appropriate counterfactuals and improve the precision of treatment effect estimation in staggered settings. I also compare and discuss their methodological differences.

The use of multiple methodologies reflects both practical and theoretical motivations. OLS offers initial correlational evidence and justifies deeper analysis. TWFE helps illustrate dynamics and potential anticipation effects but assumes homogeneous treatment impacts. The advanced DID estimators are better suited for identification in staggered treatment settings and thus form my main empirical results.

In a separate analysis, I examine the potential effect of academy creation on the quality of universities. This is motivated by historical claims that early academies played a role in raising standards in higher education. Section 2.6 analyzes university quality as a distinct outcome, shedding light on possible institutional complementarities and broader educational dynamics.

Finally, in Section 2.7, I conduct a set of additional analyses. These include explicit tests of the SUTVA assumption using spatially defined neighborhood treatment variables, sensitivity checks that exclude individual sample units, and an examination of possible spillover effects to assess whether the observed impacts extend beyond the treated cities themselves.

Taken together, this multi-method empirical approach—ranging from exploratory OLS to dynamic and advanced DID frameworks—aims to provide a robust and comprehensive understanding of the relationship between academy creation and long-run urban development in early modern Europe.

2.5 Results

Appendix Table 2.8 reports panel fixed effects regressions using Ordinary Least Squares (OLS) estimator. The dependent variable is the natural logarithm of city population, and the main independent variable is the presence of an academy. The regressions also control for the presence of a university, allowing for a comparison between different types of higher education institutions across European cities from 1500 to 1900. My preferred specification includes city fixed effects and country-by-period fixed effects, which together account for time-invariant city characteristics and broader national trends over time.[22] I do not include additional time-varying controls because there are no such data for my sample of cities. As discussed in Appendix 2.C.2, I show that the inclusion of city and country-by-period fixed effects already captures the most relevant sources of variation, mitigating omitted variable bias. This is further supported by analyses using a subset of larger cities for which more detailed time-varying data are available, following Bosker, Buringh, and Van Zanden (2013). The results in Appendix Table 2.8 show that the coefficient on the presence of an academy is positive and statistically significant across all specifications. Moreover, its magnitude is consistently larger than the corresponding coefficient for universities, suggesting that academies may have played an especially important role in driving urban growth. Appendix Table 2.9 explores heterogeneity by field of study, based on the categorization introduced in Section 2.3. This analysis provides initial evidence—later confirmed in the main results—that the positive effects are primarily driven by scientific academies.

To reduce concerns about endogeneity, I estimate a dynamic two-way fixed effects (TWFE) model and present the results using a panel event-study framework (Bhalotra et al., 2023; Clarke & Tapia-Schythe, 2021; Jacobson, LaLonde, & Sullivan, 1993).[23] [24] This strategy exploits variation in the timing of academy

[22]Country borders are defined as of the year 2000.

[23]As discussed in the introduction, this approach only partially addresses issues of reverse causality. While the DID framework helps mitigate endogeneity arising from the non-random location of educational institutions, it cannot fully rule out the influence of unobserved factors that may affect both the establishment of academies and urban growth. Based on the historical context, however, I assume that most academies were founded and managed by enlightened scholars, whose decisions were largely independent from the local economic or demographic conditions.

[24]Appendix 2.C.4 reports the static TWFE estimates.

creation across cities, allowing me to estimate the dynamic effects of treatment over time. I classify cities into two groups: the treated group consists of those that experienced the creation of at least one academy, while the control group includes cities without any academy throughout the sample period. The outcome variable is the growth rate of the natural logarithm of population, denoted as $\Delta \ln POP_{ct}$, where c indexes cities and t time periods. The event-study specification is given by:

$$\begin{aligned}\Delta lnPOP_{ct} = \beta_0 + & \sum_{l=2}^{5} \beta_l{}^{lead} EVENT_c \times \mathbb{1}\{lead_t = l\} + \\ & \sum_{k=0}^{3} \beta_k{}^{lag} EVENT_c \times \mathbb{1}\{lag_t = k\} + \\ & \mu_c + \psi_s \lambda_t + \epsilon_{ct} \end{aligned} \tag{2.1}$$

In Equation 2.1, μ_c denotes city fixed effects, and $\psi_s \lambda_t$ represents country-by-period fixed effects to account for broader national shocks and trends. The terms $\beta_l{}^{lead}$ and $\beta_k{}^{lag}$ capture, respectively, the effects in the periods leading up to and following the creation of an academy. The first lead is omitted and serves as the reference category. No additional time-varying controls are included, consistent with the fixed effects strategy discussed in Appendix 2.C.2.

By construction, this specification assumes absorbing treatment states: once a city creates an academy, the institution is assumed to remain open throughout the remainder of the sample period. As a result, the analysis estimates the intention-to-treat (ITT) effect of academy creation on urban growth between 1500 and 1900 in cities where an academy was created before 1800. This assumption of absorbing states is consistently applied across all empirical specifications, including those presented in Sections 2.5.1 and Appendix 2.C.5.[25]

2.5.1 Main findings on city population

Recent advances in the difference-in-differences (DID) literature have shown that classical dynamic two-way fixed effects (TWFE) estimators can be biased when treatment occurs at different times across units, especially in the presence of heterogeneous effects (Goodman-Bacon, 2021; Roth et al., 2023). In my context, it is reasonable to expect that the effect of founding an academy in

[25]This assumption is necessary due to data limitations. While many academies in my sample were founded in the 18th century, I do not observe precise closure dates after 1800. Collecting such data would require accounting for the disruptions caused by major historical events, such as the Napoleonic Wars, which led to temporary closures and later reopenings. In most cases, these interruptions lasted less than 50 years, making them difficult to track accurately in the panel structure of the analysis.

Oxford in 1651 differs from founding one in Turin in 1757. As the dynamic TWFE estimates in Appendix 2.C.3 suggest, heterogeneity exists not only across cities but also across time.

In simple settings—such as a 2x2 DID framework with only one treatment period and no staggered adoption—heterogeneity over time is not a concern, and negative weighting does not arise. I present these simple event studies in Appendix 2.C.5, where I estimate dynamic TWFE regressions using only one cohort of treated units at a time. The results confirm that post-treatment effects vary by cohort, supporting the presence of treatment effect heterogeneity.

In more complex staggered adoption settings, dynamic TWFE estimates aggregate 2x2 DID comparisons across all groups and time periods using implicit weights. These weights can be negative, particularly when comparing earlier- and later-treated units, leading to potentially misleading estimates (Goodman-Bacon, 2021; Jakiela, 2021). To assess this in my case, I compute the weight decomposition for the $ACADEMY$ treatment and find that the proportion of negative weights is negligible. Nonetheless, given the evidence of heterogeneity, using estimators robust to staggered treatment timing remains crucial to ensure valid inference.

To address this issue, I adopt three heterogeneity-robust estimators: those proposed by Sun and Abraham (2021), De Chaisemartin and d'Haultfoeuille (2024), and Callaway and Sant'Anna (2021). These estimators improve upon TWFE by constructing counterfactuals more carefully and avoiding inappropriate weighting, thus producing more reliable average treatment effect estimates when treatment timing and effects vary across units and periods (De Chaisemartin & D'haultfœuille, 2023).

My primary interest lies in estimating the dynamic effects of academy creation—specifically, the average impact on city population in each period following the treatment. I define cohorts as groups of cities that established an academy in the same period and aim to estimate a separate treatment effect for each cohort-period pair.

My main results rely on the interaction-weighted (IW) estimator proposed by Sun and Abraham (2021). This regression-based approach estimates a weighted average of cohort-specific average treatment effects on the treated (CATT) for each relative time period. The method interacts indicators for relative period l with indicators for cohort e, yielding interpretable $\text{CATT}_{e,l}$ coefficients. The IW estimator then averages these using non-negative, data-driven weights—computed as the proportion of treated units in each cohort-period combination—to produce consistent estimates of the overall dynamic effects.

In the event-study implementation, cities that never hosted an academy serve as the control group. As said earlier, I include outcomes through 1900 to ensure sufficient post-treatment periods, relying on the assumption of absorbing treatment states (i.e., once treated, always treated). This allows for an intention-to-treat (ITT) interpretation of the effects. However, including the nineteenth century introduces a small number of outlier observations—cities with population growth rates exceeding 200% or dropping below -100%. To obtain more conservative estimates, I exclude these outliers from the event-study analysis.[26] After this exclusion, the final sample used for the main event-study analysis includes 2,023 observations, covering 149 universities and 81 academies. Cities such as Saint Petersburg and Arras, which hosted both universities and academies and fall within the outlier range, are excluded from this part of the analysis.

Figure 2.3 presents the average treatment effect of academy creation between 1500 and 1800, estimated using the interaction-weighted estimator by Sun and Abraham (2021). The pre-trends assumption is jointly satisfied, supporting the validity of the identification strategy. Post-treatment estimates reveal that, beginning around 100 years after the creation of an academy, treated cities experience population growth approximately 9–15% higher than untreated cities (pre-treatment stats: $\mu = 0.2$, $sd = 0.3$).

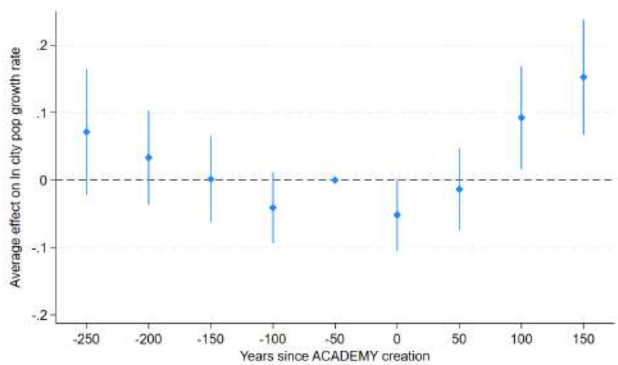


FIGURE 2.3: Academy event.

Effect of academy establishment between 1500 and 1900 on population growth estimated using the method of Sun and Abraham (2021). The control group includes cities that never established an academy.

Note: The dependent variable is the logarithmic city population growth rate. The sample includes 2020 clusters. The adjusted R^2 of the model is 0.299.

[26]Out of 73 outliers, only a few occurred before 1800. Examples include: Cacares (ESP), Chaves (PRT), Kronslot (RUS), Rochefort (FRA), Saint Petersburg (RUS), Valletta (MLT), and Versailles (FRA) which grew by more than 200%; and Arras (FRA) Bagnères-de-Bigorre (FRA), Bolgary (RUS), Burgos (ESP), Cartagena (ESP), Lucera (ITA), Salerno (ITA), Wolin (POL), Worms (DEU) and others, which experienced population declines exceeding 100%.

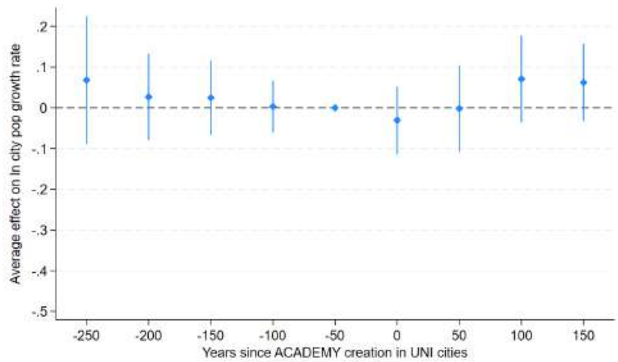


FIGURE 2.4: Academy event in university cities.

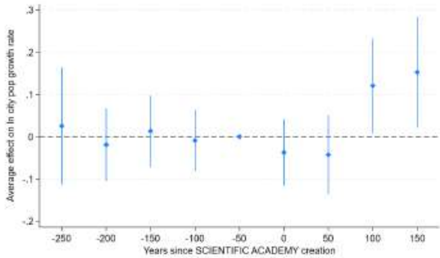
Effect of academy establishment between 1500 and 1900 on population growth in cities that hosted a university at least once, estimated using the method of Sun and Abraham (2021). The control group includes cities that never established an academy.
Note: The dependent variable is the logarithmic city population growth rate. The sample includes 142 clusters. The adjusted R^2 of the model is 0.397.

To explore the interaction between universities and academies, I restrict the sample to cities that hosted a university at least once. I then apply the same estimator to assess whether the creation of an academy in these cities produced additional effects. This reduces the sample to 149 cities. As shown in Figure 2.4, I find no significant impact of academy creation in cities that already had a university. The estimated coefficients are initially negative and very close to zero, suggesting a potentially negative or null marginal effect of adding an academy in a city with a traditional university. After 100 years, some coefficients become slightly positive, but none are statistically different from zero.

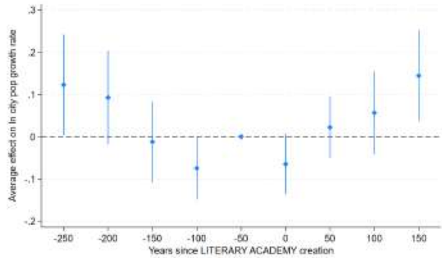
I further disaggregate the $ACADEMY$ event by examining academy characteristics: field of study, longevity, and size. Figure 2.5 shows results for scientific, literary, long-lasting, and large academies.[27] Long-lasting academies display a pattern similar to the general case, while large academies show weaker and less consistent effects.

Field of study appears to be the most important factor. Figure 2.5b shows estimates for literary academies. While a slight downward pre-trend is visible, it is not statistically significant when considered jointly (see Appendix 2.D.1). The first post-treatment coefficient is negative but quickly turns slightly positive—though statistically insignificant—up to 100 years after academy creation.

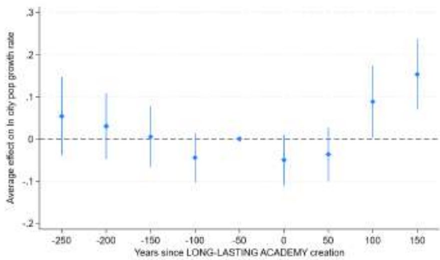
[27] Long-lasting academies are those that operated for more than 30 years; large academies are those with more than 30 members.



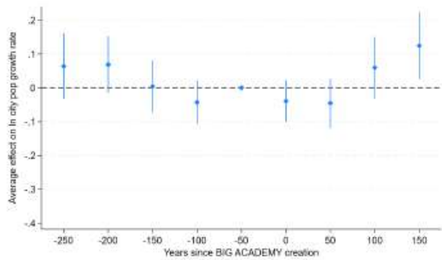
(A)



(B)



(C)



(D)

FIGURE 2.5: Academy event by field, size, length.

Effect of creating (a) a scientific academy, (b) a literary academy, (c) a long-lasting academy (with more than 30 years of activity), and (d) a big academy (with more than 30 members) between 1500 and 1900 estimated using the method of Sun and Abraham (2021). The control group includes cities that never established an academy. *Note:* The dependent variable is the logarithmic city population growth rate. The sample includes 2020 clusters. The adjusted R^2 of the model is (a) 0.298, (b) 0.298, (c) 0.299, (d) 0.299.

Interestingly, after 150 years, literary academies appear to be associated with a 14% increase in city population growth relative to cities without literary academies. However, these results are not confirmed by the alternative estimator from De Chaisemartin and d'Haultfoeuille (2024) (see Appendix 2.D.2), suggesting they should be interpreted with caution.

In contrast, scientific academies consistently show strong and statistically meaningful effects. As shown in Figure 2.5a, cities in which more than 50% of academy members were involved in science, applied science, or medicine experienced significantly higher population growth. These cities grew 12% faster 100 years after academy creation (p-value: 0.079), with the effect persisting over time and reaching 15.3% after 150 years (pre-treatment stats: $\mu = 0.20$, $sd = 0.31$).

These findings align with the OLS and dynamic TWFE estimates, the observed positive effects of academies on urban growth are primarily driven by scientific institutions. Literary academies do not appear to produce robust effects on population growth. These results highlight the importance of tracking micro-level variation in institutional characteristics—particularly the fields of study emphasized—to evaluate the role of human capital formation prior to the Industrial Revolution. This conclusion is further supported by the findings in Section 2.6, where I explore the link between academy creation and university quality.

Finally, as robustness checks, I replicate the analysis using the DID estimators developed by Callaway and Sant'Anna (2021) and De Chaisemartin and d'Haultfoeuille (2024). The results are consistent with those obtained using the interaction-weighted estimator. Full details are provided in Appendix 2.C.5.

2.6 Quality of Universities

In this section, I examine whether the establishment of academies influenced the quality of universities over time. Using the same period analyzed in Section 2.5.1, I apply an identical DID identification strategy to estimate the effect of academies on this alternative outcome.

While this analysis does not fully eliminate concerns of endogeneity, it offers important insights into the potential link between the rise of academies and the modernization of universities. In particular, it sheds light on the mechanisms by which academies may have contributed to the emergence of professionally oriented, innovation-driven universities in Europe by the 19th century.

University quality is measured using the approach in de la Croix et al. (2023), based on the aggregated quality of the top five professors who taught at a university in the preceding 25 years. I calculate the average university quality for each 50-year interval up to 1800. This allows me to replicate the DID design used in the population growth analysis, applying my preferred estimator from Sun and Abraham (2021). Two lags are included in the event structure to maintain balance timing across cohorts, as quality data is unavailable after 1800. Including more lags would introduce estimation noise, particularly since the first academy—Accademia degli Investiganti—was founded in 1650 and would be the only one considered for the estimation of the third (and last) lag.

Figure 2.6 presents the baseline results. Pre-treatment trends are flat, indicating no anticipation effects. However, I find no statistically significant change in university quality following the creation of an academy. That is, establishing an academy in a city with an existing university does not lead to improvements in university quality compared to similar cities without academies.

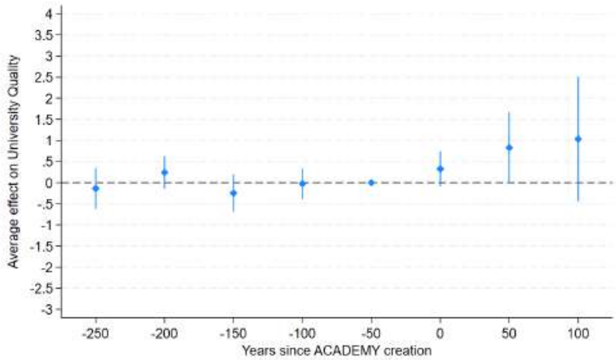


FIGURE 2.6: Academy event on university quality.

Effect of academy establishment between 1500 and 1900 on university quality in cities that hosted a university at least once, estimated using the method of Sun and Abraham (2021). The control group includes cities that never established an academy.
Note: The dependent variable is the 50-year average of university quality. The sample includes 142 clusters. The adjusted R^2 is 0.744.

This changes when I focus on scientific academies. As shown in Figure 2.7a, the creation of a scientific academy significantly improves university quality after 50 years. The estimated effect is 1.45 (p-value: 0.06), which corresponds to an average increase of over 44% relative to pre-treatment levels ($\mu = 3.27$, $sd = 2.35$). This suggests that scientific academies played a meaningful role in fostering higher academic standards within universities.

By contrast, literary academies show no such effect. As shown in Figure 2.7b, while the pre-trend is stable, the post-treatment coefficients are negative and statistically insignificant. This aligns with findings in the main analysis: literary academies do not appear to enhance either urban growth or university quality.

Finally, I examine whether academy size or longevity makes a difference. Figures 2.7c and 2.7d show that neither long-lasting nor large academies have a meaningful impact on university quality. These characteristics, while potentially reflective of institutional strength, do not translate into measurable improvements in academic standards.

Taken together, these results suggest that scientific orientation, rather than institutional scale or survival, is the key channel through which academies influenced universities. The presence of a scientific academy appears to contribute directly to the development of higher-quality institutions of learning—highlighting the importance of intellectual specialization and research orientation in shaping the evolution of European universities.

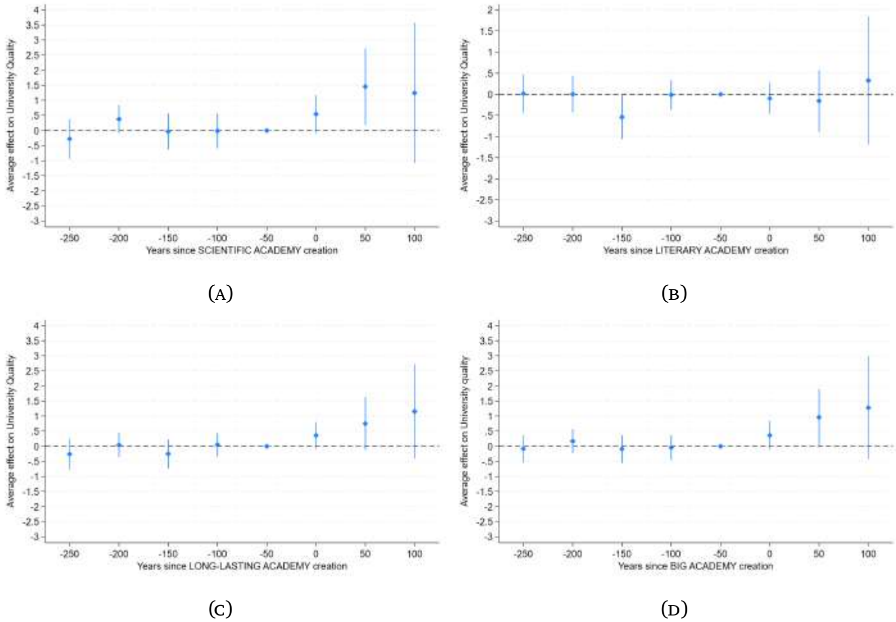


FIGURE 2.7: Academy event by field, size, and length on university quality.

Effect of creating (a) a scientific academy, (b) a literary academy, (c) a long-lasting academy (with more than 30 years of activity), and (d) a big academy (with more than 30 members) between 1500 and 1800 on university quality estimated using the method of Sun and Abraham (2021). The control group includes cities that never established an academy.
Note: The dependent variable is the 50-year average of university quality. The sample includes 142 clusters. The adjusted R^2 is (a) 0.742, (b) 0.736, (c) 0.744, (d) 0.744.

2.7 Additional analyses

This section provides additional analyses to assess the robustness and specificity of the main findings. In Section 2.7.1, I conduct leave-one-out tests by excluding key cities—such as London and Paris—and major countries that host a large number of innovative academies, including France, Italy, Germany, and Great Britain. I also test the effect of excluding Spain, a country with only one academy, to address concerns that its limited variation may disproportionately affect the construction of counterfactuals.

In Section 2.7.2 I explore the potential for spillover effects from nearby cities. The creation of an academy in one city could be influenced not only by the characteristics of that specific urban area but also by those of nearby cities, which could bias the DID estimator. I show that this is not a concern, demonstrating the coefficients for the unbiased local effect.

Finally, Section 2.7.3 presents spillover results from “donut” regressions, which explicitly test for effects extending beyond the treated city itself.

2.7.1 Sensitivity analyses: leave-one-out

Maintaining a European perspective provides a broad and generalizable framework, but it also raises the possibility that individual units—especially prominent cities or countries—might disproportionately influence the results. To address this concern, I conduct a series of leave-one-out sensitivity tests. These exclude major cities such as London and Paris, as well as countries with a high concentration of academies: France, Italy, Germany, and Great Britain. I also test the effect of excluding Spain, which contributes only one academy but many untreated observations.

London and Paris. Excluding London or Paris has no meaningful effect on the results. The estimated coefficients from the event studies—across all types of academies—remain virtually unchanged. In fact, their magnitudes slightly increase, without affecting significance. These findings indicate that the main effects are not driven solely by the most prominent scientific centers and help alleviate concerns about endogeneity.

The broader European dataset allows for the exclusion of entire countries without compromising statistical power. I begin with France, which hosts the largest share of academies (29 of 81) and universities (33 of 149).

France. Excluding France alters some dynamics. In the main specification, the parallel trends appear stronger, but there is a small, negative, and marginally significant effect immediately after academy creation. However, this is offset within the following century, where cities grow 13% faster than those without academies. The positive effect at 100 years is larger than in the baseline but does not persist.[28] This pattern may reflect France’s unique revolutionary history. In 1793, the National Convention abolished all academies linked to the *Ancien Régime*, confiscating their assets and disbanding their networks (Taillefer, 1984, p.241).[29] Removing France therefore eliminates these late-period shocks, which could bias long-run estimates downward.

[28] See Appendix Figure 2.28.

[29] The second article of the Decree on August 8, 1793 mandated the closure of academies and the confiscation of all their materials, including “books, manuscripts, medals, machines, tables, and other objects," which were placed in storage. The buildings were also seized and sold as national property in the subsequent years (Taillefer, 1984, p.241).

For other events—such as academy effects in cities that ever hosted a university, or impacts on university quality—excluding France yields results consistent with the benchmark.[30]

However, results by academy type are more sensitive. Excluding France increases the estimated long-run effect on population growth of scientific academies (16% after 100 years), though the effect again fades over time. University quality estimates remain similar in magnitude and sign, but drop in significance—likely due to reduced statistical power.[31]

For literary academies, excluding France weakens the pre-trends and reveals a small, temporary negative effect immediately after academy creation, followed by coefficients close to zero. These results, though not robust, echo theories of rent-seeking behavior (Murphy, Shleifer, & Vishny, 1991) and recent literature linking slower growth to religious-legal scholarship and traditional education systems (Curtis & de la Croix, 2023b; Squicciarini, 2020).[32]

Excluding France also anticipates the positive effect of long-lasting academies on population growth (though not on university quality), while results for large academies lose significance.[33]

Italy. Italy's exclusion does not alter the baseline event on population growth and actually strengthens the estimated effects on university quality. The statistical significance of scientific academies weakens slightly at 100 years but becomes stronger at 150 years. Literary academies remain insignificant, and pre-trends are slightly worse.[34]

Germany. Removing Germany delays the observed positive impact of academies by 50 years—likely because German academies were, on average, established later (around 1755, compared to 1741 in the full sample).[35] The most notable change is a drop in the significance of the scientific academy effect, despite Germany hosting only six such academies. This suggests those few institutions play an outsized role.[36]

Great Britain. Excluding Great Britain does not substantially affect the main results. The academy event shows a small, short-term negative effect, which is reversed after 100 years. Subgroup analyses by field, size, or duration remain consistent. However, scientific academies fall just outside standard

[30] Appendix Figures 2.28c and 2.28b.
[31] France includes 11 of 43 scientific academies. See Figures 2.29a and 2.29b.
[32] Figures 2.29c and 2.29d.
[33] Figures 2.29e, 2.29f, 2.29g, and 2.29h.
[34] Appendix Figures 2.30a to 2.31c.
[35] Although the difference is only 15 years, it aligns with the cut-off point for sample periods, shifting the results by 50 years.
[36] Figures 2.33a and 2.33b.

significance thresholds in the university quality regressions, again suggesting a loss of power.[37]

Spain. Spain hosts only one academy (Barcelona, 1764) and 20 universities. While the treatment group is largely unaffected by Spain's exclusion, the untreated pool changes meaningfully. Despite this, the main results remain robust. The only difference is a delay in the positive effects of scientific academies to 150 years after creation. Literary academies show slightly better pre-trends and continue to exhibit a long-run positive effect.[38] Estimates on university quality are also unchanged. This suggests that even in the absence of Spain, the analysis retains sufficient variation and the core findings remain valid.

2.7.2 Local effects

So far, the event studies in this paper have assumed that the effects of academy creation are purely local. However, hosting cities may themselves be influenced by nearby cities through various channels—such as shared labor markets, intellectual exchange, or regional development patterns. In such cases, the Stable Unit Treatment Value Assumption (SUTVA) may be violated. SUTVA requires that treatment effects depend only on whether a city is treated, not on the treatment status of its neighbors. To address this, a common strategy is to exclude nearby cities to account for potential "inbound" spatial spillovers—i.e., cases where neighboring cities affect the treated unit itself (Butts, 2021). This exclusion refines the construction of the counterfactual, ensuring that nearby cities which could bias the estimate of the treated city's outcome are removed. If all such potentially influential cities are excluded, the resulting estimates can be interpreted as unbiased local effects of academy creation.

Following the methodology of Johnson, Thomas, and Taylor (2023), who explore spatial spillovers in a historical context similar to this study, I exclude cities located within 50, 100, and 150 km of an academy, while keeping the hosting city in the sample.[39] If the results remain robust under these exclusions, the estimated effects can be interpreted as unbiased local impacts of academy creation.

When cities within 50 km of any academy are excluded, the main findings on population growth remain largely intact. While significance declines—expected due to the smaller sample (1,614 cities, down from 2,023)—the core pattern persists.

Excluding cities within 100 and 150 km slightly reduces the magnitude of the positive effects observed 100 years after academy creation. At this horizon,

[37] Figure 2.35b.
[38] Figures 2.37a and 2.37c.
[39] I do not extend the exclusion radius beyond 150 km to preserve statistical power.

the estimates are no longer statistically relevant. However, the long-run effects (after 150 years) remain positive and highly significant. For university quality, the results also persist.[40]

I repeat the analysis for scientific, literary, long-lasting, and large academies, too, excluding cities near each respective type. For scientific academies, the results are particularly robust. Excluding nearby cities within 50, 100, or 150 km does not materially alter the findings. The event study reveals a 16.9% increase in population growth rate 150 years after creation, alongside a substantial 86% average improvement in university quality within the first 50 years (relative to a pre-treatment mean of 1.70 and standard deviation of 2.18).[41] Literary academies also show consistent trends when excluding cities within the first 50km. However, when excluding cities within 100 km or more, the positive long-run effect disappears, and a slightly significant short-term negative impact emerges, suggesting again that the results for literary academies should be interpreted with caution.[42] For long-lasting and large academies, the spillover adjustments do not change the main patterns. The estimates remain positive but become slightly less statistically significant as the exclusion radius increases to 150 km.[43]

2.7.3 Spillover effects

In this section, I investigate "outbound" spatial spillover effects from academy creation, complementing the previous analysis of unbiased local effects. While the earlier analysis focused on whether nearby cities influence the treated city, I now assess whether the creation of an academy affects neighboring cities themselves. This follows standard spatial analysis techniques (Butts, 2021; Keller & Shiue, 2021). Specifically, I implement a "donut regression" approach: I construct concentric zones—donuts—around academy-hosting cities, using cities within these rings as the treated group, and those further away as controls. The hosting city is excluded to isolate spillovers. Due to constraints on statistical power, I limit the analysis to cities within 75 km, split into three bands: 0–25 km, 25–50 km, and 50–75 km.[44]

The results show a short-term positive spillover within the 0–25 km band: nearby cities experience a temporary increase in population growth for about 50 years following academy creation. However, this effect fades over time. In contrast, the hosting cities maintain long-run benefits. This may suggest that

[40] See Figure 2.38.
[41] Figure 2.40.
[42] Figure 2.41.
[43] Figures 2.42 and 2.43.
[44] I explored spillovers up to 150 km, but results beyond 100 km were not statistically significant due to sample limitations.

nearby cities benefit from initial knowledge diffusion or reputational spillovers without incurring the institutional costs, while only the hosting city captures persistent gains. Similar patterns appear—though with weaker significance—in the 25–50 km band and disappear entirely beyond 50 km. When disaggregating by academy type, scientific academies exhibit a positive but only slightly statistically significant short-run spillover in the 0–25 km ring, which does not extend beyond. This result is somewhat surprising, as one might expect scientific institutions to exert broader regional influence. By contrast, literary, long-lasting, and large academies do not exhibit measurable outbound spillovers at any distance, with coefficients close to zero throughout.[45]

2.8 Conclusions

In this paper, I show the long-term role of Early Modern Academies in the economic growth of European cities. At the heart of this study is a unique dataset of academicians active between 1500 and 1800, combined with a robust DID design to analyze the establishment and evolution of these educational institutions. Evaluating economic growth through population growth rates, I find that academies had a positive effect—particularly science-focused academies, which exhibited a stronger and more persistent impact.

When examining their impact on local university quality, I observe a strong positive effect from the establishment of purely scientific academies. Similarly, only scientific academies produced short-term positive spillover effects beyond the local area. Finally, the analysis of historical interactions between academies and universities opens the door to further research on the origins of the modern, professionally oriented educational system.

These results contribute to our understanding of the complex relationship between human capital, science, and economic dynamics in historical contexts. I emphasize the importance of experimental approaches and practical research in driving economic growth—even in pre-industrial times—in line with the literature on "useful knowledge" (Mokyr, 2005b). At the same time, I acknowledge the Eurocentric focus of this history of science: my analysis centers on European scientific institutions, methods, and scholars, as the West was the first to experience the Industrial Revolution. Nevertheless, future research could extend this analysis to include contributions from other civilizations—particularly Arabic and Chinese traditions—and their connections to the development of the modern scientific method (Bala, 2006; H. F. Cohen, 1994; Needham, 1964).

[45] See Appendix 2.E.3 for full results.

Acknowledgments

I am grateful to David de la Croix, Joseph F. Gomes, Luca Pensieroso, Joel Mokyr, Walker W. Hanlon, Ralf Meisenzahl, and Paula Gobbi for guidance. I thank the Research Assistants working on the ERC project 883033 and on the PDR A2.11903.007-F for the help with the data collection.

I have benefited from discussions with Larbi Alaoui, Mike Andrews, Sascha O. Becker, Enrico Berkes, Christine Binzel, Raouf Boucekkine, Mario Carillo, Matteo Cervellati, Giorgio Chiovelli, Damian Clarke, Matthew Curtis, Marcos Demetry, Claude Diebolt, Oded Galor, Victor Gay, Cecilia Garcia-Peñalosa, Antoine Germain, Michela Giorcelli, Johannes Johnen, Noel Johnson, Wolfgang Keller, Jean-Francois Maystadt, Stelios Michalopoulos, Enrico Moretti, Peter Nencka, Sebastian Ottinger, Nuno Palma, Luigi Pascali, Giovanni Peri, Mathilde Portois, Hillel Rapoport, Gérard Roland, Lukas Rosenberger, Jared Rubin, Jonathan Schulz, Mara Squicciarini, Uwe Sunde, Felipe Valencia Caicedo, Nico Voigtländer, David Weil, and Giulio Zanella. I also thank the conference participants to the Winter Econometric Society 2023, Clio 2024, Economic History Association 2024, Economic History Society 2025, World Economic History Congress 2025.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, under grant agreement No 883033 "Did elite human capital trigger the rise of the West? Insights from a new database of European scholars". This work was also supported by the Fonds de la Recherche Scientifique-FNRS under Grant n° A2.11903.007-F "Human capital and the rise of the West: the key role of scientific academies".

2.9 Bibliography

Academiae Electoralis Scientiarum et Elegantiorum Literarum Theodoro-Palatina. (1766). *Historia et commentationes academiae electoralis scientiarum et elegantiorum literarum theodoro-palatinae. volumen 1* [From the collections of Ghent University. Public Domain.]. Mannheimii: Typis Academicis. https://books.google.com/books?id=iVUVAAAAQAAJ

Académie de Béziers. (2024). Tricentenaire Academie Beziers 1723 - 2023 [Consulted in April 2024]. https://tricentenaire-beziers.fr/historique/

Académie De Cherbourg. (2024). Historique [Consulted in May 2024]. https://academiedecherbourg.wordpress.com/la-societe-nationale-academique-de-cherbourg/

Académie des sciences, des lettres et des arts d'Amiens. (1901). *Tome XLVIII. Mémoires de l'Académie des sciences, agriculture, commerce, belles-lettres et arts du département de la Somme*. Imprimerie Yvert et Tellier.

Académie des Sciences, Lettres et Arts d'Arras. (2024). Histoire de l'Académie d'Arras [Consulted in March 2024]. https://academie-arras.fr/histoire

Académie des Sciences Lettres et Arts de Marseille. (2024). Histoire de l'Academie depuis 1726 [Consulted in April 2024]. https://www.academie-sla-marseille.fr/pages/l-academie-de-marseille/histoire-de-l-academie-depuis-1726-1720124475.html

Académie Royale de Marine. (2024). 250 ans d'histoire [Consulted in March 2024]. https://www.academiedemarine.com/250ans.php

Académie Royale des sciences. (2024). Histoire de l'Académie des sciences [Consulted in February 2023]. https://www.academie-sciences.fr/fr/Histoire-de-l-Academie-des-sciences/histoire-de-l-academie-des-sciences.html

Académie Royale des Sciences, belles-lettres et arts de Lyon. (2024). Histoire de l'Académie [Consulted in April 2024]. https://academie-sbla-lyon.fr/Academie/histoire

Accademia delle Scienze di Torino. (2023). *Oggi si chiamano brevetti, nel passato privilegi industriali* [Consulted in June 2023]. https://accademiadellescienze.it/brevetti

Accademia Nazionale di Scienze, Lettere e Arti di Modena. (2023). I soci storici [Consulted in February 2023]. https://www.accademiasla-mo.it/it

Accademia Peloritana dei Pericolanti. (2024). Storia dell'Accademia [Consulted in March 2024]. https://www.accademiapeloritana.it/storia.html

Accademia Roveretana degli Agiati di Scienze, Lettere ed Arti. (2024). Storia [Consulted in March 2024]. https://agiati.org/storia

Allombert, F. (1899). *Annales de la Societe d'Emulation et Agriculture (Lettre, Science, et arts) de l'Ain* [Bourg-en-Bresse].

Applebaum, W. (2000). *Encyclopedia of the scientific revolution: from Copernicus to Newton*. Routledge.

Ashraf, Q., & Galor, O. (2011). Dynamics and stagnation in the Malthusian epoch. *American Economic Review*, *101*(5), 2003–2041.

Bairoch, P., Batou, J., & Pierre, C. (1988). *Population des villes européennes de 800 à 1850: Banque de données et analyse sommaire des résultats (la)*. Librairie Droz.

Bala, A. (2006). The Eurocentric History of Science. In *The Dialogue of Civilizations in the Birth of Modern Science* (pp. 21–32). Palgrave Macmillan US.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The quarterly journal of economics*, *106*(2), 407–443.

Barro, R. J. (2001). Human capital and growth. *American economic review*, *91*(2), 12–17.

Bayerische Akademie der Wissenschaften. (2024). Chronik der Akademie [Consulted in May 2024]. https://badw.de/geschichte/chronik.html

Becker, S. O., & Woessmann, L. (2009). Was Weber wrong? A human capital theory of Protestant economic history. *The quarterly journal of economics*, *124*(2), 531–596.

Benos, N. (, Conti, M., Papazoglou, M., & Tsoumaris, S. (2024). The Rise of Universities and Economic Development in Medieval and Early Modern Europe (700–1800). *SSRN http://dx.doi.org/10.2139/ssrn.4967505*.

Berkes, E., & Nencka, P. (2021). Knowledge access: The effects of Carnegie libraries on innovation. *Available at SSRN 3629299*.

Berry, H. F. (1915). *A history of the royal dublin society* [London]. Longmans, Green and Co.

Bhalotra, S., Clarke, D., Gomes, J., & Venkataramani, A. (2023). Maternal mortality and women's political participation. *Forthcoming at the Journal of the European Economic Association*.

Bianchi, N., & Giorcelli, M. (2020). Scientific education and innovation: From technical diplomas to university stem degrees. *Journal of the European Economic Association*, *18*(5), 2608–2646.

Bierens de Haan, J. A. (1952). *De Hollandsche Maatschappij der Wetenschappen, 1752-1952.*

Böhme-Kaßler, K. (2005). *Gemeinschaftsunternehmen Naturforschung : Modifikation und Tradition in der GNF zu Berlin 1773 - 1906* [Stuttgart]. Franz Steiner Verlag GmbH.

Bois, J.-P. (2021). *Histoire de L'Académie Royale des Belles-Lettres d'Angers, 1685-1793.* [Angers]. Le Coudray-Macouart, Saint-Léger Editions.

Boissier, G. (1907). *La Suppression des Académies en 1793* [5e période, Vol. 40, No.4, pages 721-751]. Revue des Deux Mondes (1829-1971).

Bosker, M., Buringh, E., & Van Zanden, J. L. (2013). From Baghdad to London: Unraveling urban development in Europe, the Middle East, and North Africa, 800–1800. *Review of Economics and Statistics*, *95*(4), 1418–1437.

Brizzi, G. P. (2001). *L'antica Università di Fermo*. Milano-Fermo, Silvana.

Buringh, E. (2021). The Population of European Cities from 700 to 2000: Social and Economic History. *Research Data Journal for the Humanities and Social Sciences*, *1*(aop), 1–18.

Butts, K. (2021). Difference-in-differences estimation with spatial spillovers. *Preprint arXiv:2105.03737*.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Campori, G. (1887). *La società filopatria di Torino*. Giornale Storico della Letteratura Italiana.

Cantoni, D., & Yuchtman, N. (2014). Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics*, *129*(2), 823–887.

Cassidy, D. C. (1985). *Meteorology in Mannheim: The Palatine Meteorological Society, 1780-1795* [Stuttgart]. Franz Steiner Verlag Wiesbaden GmbH.

Chen, Z., & Ma, C. (2022). Prioritizing Stability over Productivity: How Confucianism Facilitated Population Growth in Preindustrial China. *SSRN*.

Cinnirella, F., & Streb, J. (2017). The role of human capital and innovation in economic development: evidence from post-Malthusian Prussia. *Journal of economic growth*, *22*, 193–227.

Clarke, D., & Tapia-Schythe, K. (2021). Implementing the panel event study. *The Stata Journal*, *21*(4), 853–884.

Cohen, D., & Soto, M. (2007). Growth and human capital: Good data, good results. *Journal of economic growth*, *12*(1), 51–76.

Cohen, H. F. (1994). *The scientific revolution: A historiographical inquiry*. University of Chicago Press.

Complete list of the members and officers of the Manchester Literary and Philosophical Society, from its institution on February 28th, 1781 to April 28th, 1896 [Manchester]. (1896).

Courteault, P. (1912). *Listes des membres: 1712-1912, Actes de l'Académie nationale des sciences, belles-lettres et arts de Bordeaux, 1912, p. 327-357* [Bordeaux]. Académie nationale des sciences, belles-lettres et arts de Bordeaux.

Curtis, M., & de la Croix, D. (2023a). Measuring Human Capital: from WorldCat Identities to VIAF. *Repertorium Eruditorum Totius Europae*, *10*, 17–22.

Curtis, M., & de la Croix, D. (2023b). *Seeds of Knowledge: Premodern Scholarship, Academic Fields, and European Growth* (tech. rep.). CEPR Press Discussion Paper No. 18321.

de Colonjon, H. (1866). *Société Académique Et Patriotique De Valence, Bulletin De La Société Départementale D'archéologie Et De Statistique De La Drome*. Imprimerie De Chenevier Et Chayet.

de Pontville, M. (1997). *Histoire de l'Académie de Caen*.

De Candolle, A. (1885). *Histoire des sciences et des savants depuis deux siècles: précédée et suivie d'autres études sur des sujets scientifiques en particulier sur l'hérédité et la sélection dans l'espèce humaine*. Georg.

De Chaisemartin, C., & D'haultfœuille, X. (2023). Two-way fixed effects and differences-in-differences estimators with several treatments. *Journal of Econometrics*, *236*(2), 105480.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.

De Chaisemartin, C., & d'Haultfoeuille, X. (2022). *Difference-in-differences estimators of intertemporal treatment effects* (tech. rep.). National Bureau of Economic Research.

De Chaisemartin, C., & d'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1–45.

de la Croix, D., & Zanardello, C. (2021). Scholars and Literati at the Academy of Sciences of Turin (1757–1800). *Repertorium Eruditorum Totius Europae*, *5*, 43–51.

de la Croix, D., & Zanardello, C. (2022). Scholars at the Royal Academy of Sciences in Paris (1666-1793). *Repertorium Eruditorum Totius Europae*, *9*, 1–10. 10.14428/rete.v9i0/ASc

de la Croix, D. (2021). Scholars and Literati in European Academia before 1800. *Repertorium Eruditorum Totius Europae*, *5*, 35–41.

de la Croix, D., Docquier, F., Fabre, A., & Stelter, R. (2023). The Academic Market and the Rise of Universities in Medieval and Early Modern Europe (1000–1800). *Journal of the European Economic Association*, jvad061.

de la Croix, D., & Doraghi, M. (2021). Scholars and literati at the Saint Petersburg Academy of Sciences (1724–1800). *Repertorium eruditorum totius Europae*, *5*, 17–26.

de la Croix, D., Eisfeld, T., & Ganterer, M. (2021). Scholars and literati at the royal prussian academy of sciences (1700–1800). *Repertorium eruditorum totius Europae*, *3*, 1–9.

Defrasne, J., Maurat, J.-P., Lordereau, M., & Lassus, F. (2002). *Deux cent cinquantième anniversaire de l'académie de Besançon* [Besançon].

des Barres, C. (1851). *Discours de M. le Président sur l'histoire des anciennes Sociétés littéraires et scientifiques à Auxerre. Bulletin de la société des sciences historiques et naturelles de l'Yonne.* [Auxerre]. Perriquet.

Deutsche Akademie der Naturforscher Leopoldina. (2024). History of the Academy [Consulted in April 2024]. https://www.leopoldina.org/en/about-us/about-the-leopoldina/history/the-history-of-the-leopoldina/

Dittmar, J. (2019). The economic origins of modern science: Technology, institutions, and markets. *London School of Economics*.

Dittmar, J., & Meisenzahl, R. (2022). The research university, invention, and industry: Evidence from german history.

Dulieu, L. (1975). *La médecine à Montpellier, Tome 1, Le Moyen-Age*. Les presses universelles.

Emerson, R. L. (1981). *The Philosiphical Society of Edinburgh 1748-1768* [Cambridge]. Cambridge Unversity Press.

Ercolani, G. B. (1881). *Accademia delle scienze dello Istituto di Bologna dalla sua origine a tutto il MDCCCLXXX* [Bologna]. Nicola Zanichelli.

Flouret, J. (2009). L'académie de La Rochelle, ses activités, ses membres, de 1732 à nos jours [Bordeaux]. *Actes des congrès nationaux des sociétés historiques et scientifiques*.

Forestié, E. (1888). *La société Littéraire et l'ancienne Académie de Montauban*.

Frijhoff, W. (1996). Patterns. In H. d. Ridder-Symoens (Ed.), *A history of the University in Europe. vol. ii: Universities in Early Modern Europe (1500–1800)*. Cambridge University Press.

Gage, A. (1938). *A history of the Linnean Society of London* [London]. Taylor; Francis.

Gertz, O. (1940). *Kungl. Fysiografiska Sällskapet I Lund 1772-1940*. Häkan Ohlssons Boktryckeri.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Gordin, M. (2000). The Importation of Being Earnest - The Early St. Petersburg Academy of Sciences. *The University of Chicago Press on behalf of The History of Science Society*, *91:1*, 1–31.

Gosseaume, M. (1985). *Précis Analytique Des Travaux De L'académie Royale Des Sciences, Belles - Lettres Et Arts De Rouen* [Rouen]. P.Periaux.

Gunther, R. T. (1925). *Early Science in Oxford - VOL IV - The Philosophical Society* [Oxford].

Hanlon, W. W. (2022). *The rise of the engineer: Inventing the professional inventor during the Industrial Revolution* (tech. rep.). National Bureau of Economic Research.

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of economic literature*, *46*(3), 607–68.

Hasquin, H. (2009). *L'académie impériale et royale de Bruxelles, ses académiciens et leurs réseaux intellectuels au XVIII siècle* [Bruxelles]. Académie Royale de Belgique.

Histoire et mémoires de la société des sciences physiques de Lausanne. Tome Second. Années 1784, 1785 & 1786. [Hombourg-ez-Monts]. (1789). Mourer.

Hollandsche Maatschappij der Wetenschappen. (2024). Geschiedenis [Consulted in April 2024]. https://khmw.nl/

Jacobson, L. S., LaLonde, R. J., & Sullivan, D. G. (1993). Earnings losses of displaced workers. *The American Economic Review*, 685–709.

Jakiela, P. (2021). Simple diagnostics for two-way fixed effects. *preprint arXiv:2103.13229*.

Johnson, N. D., Thomas, A. J., & Taylor, A. N. (2023). The impact of the black death on the adoption of the printing press. *Working Paper*.

Karlberg, G. (1977). *Royal Society of Sciences of Uppsala, Catalogue of Publications 1716-1967* [Uppsala]. Kungl. Vetenskaps-societeten i Uppsala.

Keller, W., & Shiue, C. H. (2021). *The economic consequences of the opium war* (tech. rep.). National Bureau of Economic Research.

Kiefer, J. (2004). *Bio-bibliographisches Handbuch der Akademie Gemeinnütziger Wissensch- aften zu Erfurt: 1754–2004* [Erfurt]. Akad. Gemeinnütziger Wiss. zu Erfurt.

Knowles Middleton, W. (1971). *The Experimenters: a study of the Accademia del Cimento* [Baltimore and London]. The John Hopkins Press.

Königlich-Preußische Akademie der Wissenschaften. (2024). Akademiegeschichte [Consulted in May 2024]. https://www.bbaw.de/die-akademie/akademie-historische-aspekte

Koschnick, J. (2025). *Teacher-directed scientific change: The case of the English Scientific Revolution* (tech. rep.). EHES Working Paper, No. 274.

Koschnick, J., Hornung, E., & Cinnirella, F. (2022). Flow of Ideas: Economic Societies and the Rise of Useful Knowledge.

Kostlán, A. (1996). *Societas Incognitorum - První Učená Společnost V Českých Zemích* [Prague]. Archiv Akademie věd České Republiky.

Krahnke, H. (2001). *Die Mitglieder der Akademie der Wissenschaften zu Göttingen 1751-2001* [G"ottingen]. Vandenhoeck & Ruprecht.

Kungliga Vetenskapsakademien. (2024). Akademiens historia [Consulted in May 2024]. https://www.kva.se/om-oss/akademiens-historia/

Laouenan, M., Bhargava, P., Eyméoud, J.-B., Gergaud, O., Plique, G., & Wasmer, E. (2022). A cross-verified database of notable people, 3500BC-2018AD. *Scientific Data*, *9*(290).

Lauzun, P. (1900). *La société académique d'Agen (1776-1900)* [Paris]. Alphonse Picard et Fils.

Lettres Patentes [Grenoble]. (1790). Imprimerie Royale.

Lieburg, M. J. (1985). *Het Bataafsch Genootschap der Proefondervindelijke Wijsbegeerte te Rotterdam (1769-1984). Een bibliografisch en documenterend overzicht* [Amsterdam]. Editions Rodopi.

Lomholt, A. (1950). *Det Kongelige Danske Videnskabernes Selskab 1742-1942*.

Luscombe, D. E. (1969). The School of Laon. In *The school of peter abelard: The influence of abelard's thought in the early scholastic period* (pp. 173–182). Cambridge University Press.

Maggiolo, A. (1983). *I soci dell'Accademia patavina dalla sua fondazione (1599)* [Padova]. Accademia Patavina delle Scienze Lettere ed Arti.

Maylender, M. (1930). *Storia delle accademie d'Italia* (Vol. 1-5). L. Cappelli.

McClellan, J. E. (1985). *Science Reorganized: Scientific Societies in the Eighteenth Century*. Columbia University Press.

Mège, F. (1884). *L'académie des sciences, belles lettres et arts de Clermont Ferrand, Ses origines et ses travaux* [Clermont Ferrand]. Ferdinand Thibaud.

Menu, H. (1869). *Notes historiques sur la Société Littéraire et sur la Société des Sciences Arts et Belles Lettres de Chalons sur Marne 1750-1792* [Chalons-sur-Marne]. Le Roy.

Milsand, P. (1871). *Notes et documents pour servir à l'histoire de l'Académie des sciences, arts et belles-lettres de Dijon* [Dijon]. Auguste Aubry.

Mokyr, J. (2003). The gifts of Athena: Historical origins of the knowledge economy. Princeton University Press.

Mokyr, J. (2005a). The intellectual origins of modern economic growth. *The Journal of Economic History*, *65*(2), 285–351.

Mokyr, J. (2005b). Long-term economic growth and the history of technology. In *Handbook of economic growth* (pp. 1113–1180, Vol. 1). Elsevier.

Mokyr, J. (2016). A Culture of Growth. Princeton University Press.

Mokyr, J., & Voth, H.-J. (2009). Understanding growth in early modern Europe. *Cambridge Economic History of Europe*, *1*, 7–42.

Murphy, K. M., Shleifer, A., & Vishny, R. W. (1991). The allocation of talent: Implications for growth. *The quarterly journal of economics*, *106*(2), 503–530.

Needham, J. (1964). Science and Society in East and West. *Science and Society*, *28*(4), 385–408. Retrieved July 10, 2025, from http://www.jstor.org/stable/40401068

Nelson, R. R., & Phelps, E. S. (1966). Investment in Humans, Technological Diffusion, and Economic Growth. *The American Economic Review*, *56*(1/2), 69–75.

Nicolas, M. (1854). *Histoire Litteraire de Nimes et des localites voisines qui forment actuellement le departement du Gard*. Chez Ballivet et Fabre, imprimeurs.

Nicolas, M. (1908). *Mémoires de la société d'agriculture, sciences belles-lettres et arts d'Orléan*.

Niedersächsische Akademie der Wissenschaften zu Göttingen. (2024). Über die Akademie [Consulted in April 2024]. https://adw-goe.de/ueber-uns/uebersicht/

Ó Gráda, C. (2016). Did science cause the industrial revolution? *Journal of Economic Literature*, *54*(1), 224–239.

Oberlausitzische Gesellschaft der Wissenschaften. (2024). Geschichte - Von der Gründung bis zum Ende des 18. Jahrhunderts [Consulted in April 2024]. https://www.olgdw.de/gesellschaft/geschichte/

Pahl, J.-P., & Ranke, H. (2019). *Von der höheren fachschule zur fachhochschule. entstehung und etablierung einer institution* (Vol. 28).

Pepe, L. (2008). Les milieux savants en italie dans la deuxième moitié du 18 e siècle. *Dix-huitième siècle*, (1), 211–228.

Pixton, P. B. (1998). The misfiring of german cultural leadership in the twelfth century: The evidence from the cathedral schools. *Paedagogica historica*, *34*(2), 346–363.

Programme de la Société patriotique de Hesse-Hombourg, pour l'encouragement des connoissances et des moeurs [Hombourg-ez-Monts]. (1777). Société Patriotique.

Ramankutty, N., Foley, J. A., Norman, J., & McSweeney, K. (2002). The global distribution of cultivable lands: Current patterns and sensitivity to possible climate change. *Global Ecology and biogeography*, *11*(5), 377–392.

Rasmussen, W. (1969). De tekniske skolers historie. *Årbog for dansk skolehistorie*, *3*, 7–41.

Real Academia de Ciencias Exactas, Físicas y Naturales de España. (2024). Nuestra Historia [Consulted in October 2024]. https://rac.es/sobre-nosotros/nuestra-historia/1582/

Reial Acadèmia de Ciències i Arts de Barcelona. (2024). Història de la RACAB [Consulted in April 2024]. https://www.racab.cat/historia/

Réka, L., & Tüskés, G. (2017). *Learned Societies, Freemasonry, Sciences and Literature in 18th century Hungary*. MTA Bölcsészettudományi Kutatóközpont Irodalomtudományi Intézet.

Roche, D. (1964). La diffusion des Lumières. Un exemple : l'académie de Châlons-sur-Marne. *Annales. Economies, sociétés, civilisations.*, (5), 887–922.

Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy*, *98*(5, Part 2), S71–S102.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*.

Royal Irish Academy. (2024). History of the Academy [Consulted in April 2024]. https://www.ria.ie/about/history-of-the-academy/

Rübel, E. (1947). *Geschichte der Naturforschenden Gesellschaft in Zürich auf das Jahr 1947* [Zurich]. Kommissionsverlag Gebr. Fretz AG.

Schmidt, O. (1960). *Det kongelige norske videnskabers selska: Matrikkel 1760-1960* [Trodheim]. Det Kongelige Norske Videnskabers Selskab.

Schofield, R. E. (1963). *The Lunar Society of Birmingham: A Social History of Provincial Science and Industry in Eighteenth Century England* [Oxford]. Clarendon Press.

Schumann, E. (1893). *Geschichte der Naturforschenden Gesellschaft in Danzig 1743 — 1802* [Danzig]. Commissions-verlag von Wilhelm Engelmann.

Schütte, F. (2007). Jahrzehnt der Neuordnung 1890-1901. Die Reform des technischen und allgemeinen Bildungssystems in Deutschland. Berufspädagogische Anmerk- ungen zu einem bildungshistorisch'disparaten'Forschungsfeld. *Zeitschrift für Pädagogik*, *53*(4), 544–561.

Serafinelli, M., & Tabellini, G. (2022). Creativity over time and space. *Journal of Economic Growth*, *27*(1), 1–43.

Singels. (1923). *Gedenkboek van het Provinciaal Utrechtsch genootschap voor kunsten en wetenschappen 1773-1923. Uit de geschiedenis van het P.U.G. (Provinciaal Utrechtsch Genootschap) 1773-1923* [Utrecht]. Utrecht / A. Oosthoek.

Société Royale des Sciences. (2024). Histoire résumée de l'Académie des Sciences et Lettres de Montpellier [Consulted in April 2024]. https://www.ac-sciences-lettres-montpellier.fr/academie/histoire

Spalding Gentlemen's Society. (2025). Our History [Consulted in July 2025]. https://www.sgsoc.org/our-story

Squicciarini, M. P. (2020). Devotion and development: Religiosity, education, and economic progress in nineteenth-century france. *American Economic Review*, *110*(11), 3454–3491.

Squicciarini, M. P., & Voigtländer, N. (2015). Human capital and industrialization: Evidence from the age of enlightenment. *The Quarterly Journal of Economics*, *130*(4), 1825–1883.

Stanislas, A. (2024). *Historique* [Consulted in April 2024]. https://www.academie-stanislas.org/historique-general/

Stelter, R., de la Croix, D., & Myrskylä, M. (2021). Leaders and laggards in life expectancy among European scholars from the sixteenth to the early twentieth century. *Demography*, *58*(1), 111–135.

Sturges, P. (1978). *The Membership of the Derby Philosophical Society, 1783–1802* (Vol. 4, Issue 3) [Birmingham]. Midland History.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Tabarrini, M. (1856). *Degli Studi e delle vicende della Reale Accademia dei Georgofili - nel primo secolo della sua esistenza*. M. Cellini & Co., Firenze.

Taillefer, M. (1984). *L'Académie des sciences, Inscriptions, et Belles lettres de Toulouse au XVIIe siècle*. CNRS.

Teixeira Rebelo da Silva, J. A. (2015). *A Academia Real das Ciências de Lisboa (1779-1834): ciências e hibridismo numa periferia europeia* [Thesis, Universidade de Lisboa].

The Royal Society. (2024). History of the Royal Society [Consulted in January 2024]. https://royalsociety.org/about-us/who-we-are/history/

Vallauri, T. (1875). *Storia delle università degli studi del Piemonte*. Stamperia reale di GB Paravia.

Zacek, J. F. (1968). The "Virtuosi" of Bohemia: The Royal Bohemian Society of Sciences. *East European Quarterly - Periodicals Archive Online*, *2*, 147–159.

Zanardello, C. (2022). Scholars and Literati at the University of Turin(1404 – 1800). *Repertorium Eruditorum Totius Europae*, *7*, 31–38.

Zeeuwsch Genootschap der Wetenschappen. (2024). Historie [Consulted in April 2024]. https://kzgw.nl/kzgw/historie/

2.A Additional details on historical context

2.A.1 Academies' characteristics by country

In this Section, I list the main characteristics of the official recognition, topics of study, memberships and meetings, general governance, and financing of academies by modern country:

- France

Official Recognition. Generally, French academies received official recognition from the King relatively quickly, often within 10 years of their informal foundation. For example, the *Académie des Sciences, belles lettres et arts in Besançon* received official recognition in 1752, just a few years after it was formed in 1748 (Defrasne et al., 2002). The first meeting of the academy of Nancy was held on December 28, 1750, and the patent letters arrived exactly one year later, on December 27, 1751 (Stanislas, 2024). Most academies received these patent letters, signifying both royal approval and often some form of financial support either from the Royal Court or from local authorities (e.g., lords and bishops).

Topics of Study. French academies covered a wide range of subjects, including natural sciences, humanities, arts, and practical issues related to local regions. The emphasis on applied science and knowledge for the betterment of society was particularly evident in academies like those in Châlons-sur-Marne and Cherbourg. In Châlons-sur-Marne, the academy's motto was *"L'Utilité*" (literally "The Utility", (Roche, 1964)), signalling a focus on improving the living standards of the local community. The *Société Académique* in Cherbourg organized local competitions to incentivize young researchers in hydrography (Académie De Cherbourg, 2024). The experimental approach and the influence of the "New Science" were significant factors, as is clear from the creation of the *Académie Royale des Sciences* in Paris. The latter was officially founded by the minister Colbert "to advance scientific knowledge and promote the practical application of scientific research" (de la Croix & Zanardello, 2022, p.1).

Memberships and Meetings. French academies typically had three membership categories: ordinary members, who usually resided in the city of the academy; honorary members, who were influential personalities who brought fame and reputation to the institutions without advancing knowledge themselves, and correspondents, often foreign, who did not reside permanently in the city but sent letters with their thoughts and findings to be read by the ordinary members during meetings. There was often an upper limit on the number of ordinary members, explicitly written in the *Statutes* and varying according to the academy, but there was no limit on correspondents and honorary members.[46]

[46] Sometimes honorary members were also limited to a specific number, but for this category, exceptions were applied more regularly than the rule itself.

Meetings were generally regular, often weekly or bi-weekly, with meeting days written into the *Statutes* of the academies.

Governance. Most French provincial academies had a hierarchical structure very similar to the Paris Academy, with directors, secretaries, treasurers, librarians, and often a protector (usually a high-ranking figure like the King, or a local noble or bishop).[47] The directors or presidents of the academy were elected members and held office for a certain number of months. Secretaries and librarians were also elected members but usually stayed in their roles permanently. Treasurers were often also the vice-directors, but for this role, there was more variability among the provincial academies.

Finances. Finances often came from a combination of royal support, member fees, private donations, and sometimes bequests. Some academies received annual subsidies from the King (e.g. the Paris Academy, Boissier (1907)), others only at the beginning of their activities, while others relied on private contributions to achieve some independence from the central power.

- Italy

Official Recognition. Italian academies were often established informally before receiving official recognition from local rulers or the Pope. Official recognition was sometimes delayed, with some academies operating for several decades before receiving official status. For instance, it was 26 years before the King of Savoy recognized the Academy of Sciences in Turin. While many Italian academies were recognized by local rulers, the Pope's authority was significant, particularly for institutions like the *Istituto delle Scienze* in Bologna, which relied heavily on Pope Benedict XIV's donations.

Topics of Study. Italian academies were inspired by the "New Science" and its experimental approach, particularly in the early period, thanks to the influence of scholars like Galileo Galilei and his followers. Academies such as the *Accademia del Cimento* in Florence (founded by two students of Galileo) (Knowles Middleton, 1971; Maylender, 1930) and the *Accademia degli Investiganti* in Naples (Maylender, 1930) are examples of early Italian institutions focused on experimental research. Italian academies frequently focused on natural philosophy, physics, mathematics, and astronomy, but also explored literature, history of the homeland, and practical issues like agriculture. In Florence, the Georgofili Academy worked closely with local authorities to reduce the impact of famines in 1791/1792, giving advice which proved to be beneficial and thus gaining credibility (Tabarrini, 1856).

[47] From here on, when I use the term 'hierarchical' structure, I am referring to the French-like organization, where specific figures are either elected or appointed for life or for a set term.

Memberships and Meetings. Italian academies had various membership categories, including ordinary members, honorary members, and foreign correspondents, but in a much less centralized manner than their French provincial counterparts. The *Statutes* of Italian academies rarely contain rules about the maximum number of members. Meeting frequency varied significantly depending on the academy, with some meeting only once a month.

Governance. Governance structures were diverse, ranging from the more to less democratic. Some academies were overseen by a patron, while others had elected leaders and committees. The *Accademia ducale dei Dissonanti* of Modena had in its name the word "ducale" (i.e., literally "of the Duke"), explicitly indicating the Ducal patronage and his high level of influence over academic matters (Accademia Nazionale di Scienze, Lettere e Arti di Modena, 2023).

Finances. Italian academies had various sources of funding, mainly from patronage and member fees, but also from private donations and sometimes subsidies from the local or provincial church. Royal patronage was particularly important for some institutions, as mentioned above, but its generosity varied significantly.

- Germany

Official Recognition. For German academies, official recognition did not follow any specific pattern. It could come from local rulers or patrons, with variations in the timing and type of recognition. For example, the academy of Göttingen was directly established as a "Royal Society" by King George II of Great Britain and Ireland, and Elector of Hanover (Niedersächsische Akademie der Wissenschaften zu Göttingen, 2024). The city of Mannheim hosted two academies,[48] thanks to the Elector Palatine of Bavaria, Karl Theodor (Cassidy, 1985). Not obtaining recognition often meant more independence, as in the case of the academy of Görlitz, which did not obtain any official recognition, and remained a private society from its foundation in 1779 (Oberlausitzische Gesellschaft der Wissenschaften, 2024).

Topics of Study. German academies covered a wide range of subjects, including natural sciences, humanities, and applied sciences. There was a strong emphasis on experimental research and the use of empirical data, as seen in academies like the *Gesellschaft Naturforschender Freunde* in Berlin and the *Naturforschende Gesellschaft* in Jena. The former focused on producing original research on natural history thanks to their own data collections (Böhme-Kaßler, 2005). The latter aimed to supplement university lessons with more empirical applications through their collection of instruments and their own laboratory (Böhme-Kaßler, 2005).

[48] The *Academia Electoralis Scientiarum et Elegantiorum Literarum Theodoro-Palatina* created in 1763 and the *Societas Meteorologicae Palatinae* founded in 1780.

Memberships and Meetings. German academies had a mix of ordinary members, honorary members, and sometimes foreign correspondents. Meetings were generally regular, with weekly or bi-weekly gatherings being common.

Governance. Unlike France, governance structures in Germany were diverse, ranging from more informal and democratic to more hierarchical and patron-driven.

Finances. Finances were typically derived from a combination of member fees and donations, with patron support and government subsidies both being important sources. Patronage was particularly important in Germany, and the stability of the academy was highly dependent on it.

- Great Britain

Official Recognition. British academies were mostly informal, with only the Royal Society of London and the academy of Edinburgh obtaining official recognition. The timing differed significantly between the two: the Royal Society was recognized in 1662, about two years after its creation, while Edinburgh waited 52 years to obtain its royal charter in 1783.[49]

Topics of Study. British academies were primarily devoted to natural philosophy and scientific experimentation, taking the *Royal Society* as a model. Its motto *"nullius in verba*" (which translates "into take nobody's word for it") is a clear statement of the will to use the experimental perspective to test and verify every fact and conclusion (The Royal Society, 2024). The experimental approach was central to their work, as demonstrated by institutions like the *Lunar Society* of Birmingham, which focused on applied science and its relevance to industry, with members like James Watt (a mechanical engineer who worked on steam engines), Erasmus Darwin (natural philosopher, poet, and grandfather of Charles), and Richard Lovell Edgeworth (grandfather of Francis Ysidro)[50] (Schofield, 1963).

Memberships and Meetings. British academies typically had various membership categories, including ordinary fellows, honorary members, and foreign correspondents. However, for many British academies, including the Royal Society, it is impossible to distinguish between ordinary and correspondent members, as all members are called "fellows" in their records. In our database, the category is identified only for specific years for which we know the list of foreign members as detailed in De Candolle (1885).

[49]Edinburgh presents a complex case: the 52 years are calculated from the founding of its predecessor, the *"Philosophical Society of Edinburgh"*. Although it was inactive for a period, it later resumed with only minor changes in its membership, which is why we regard both societies as a single academy, established in 1731.

[50]Francis Ysidro is considered the pioneer of utility theory with his development of indifference curves and the Edgeworth box.

Meetings were generally regular, often weekly, with specific days designated for gatherings.

Governance. Most academies had a hierarchical structure, with presidents, secretaries, treasurers, and sometimes other elected officials.

Finances. British academies had very similar finances to those in other countries, with a combination of member fees and donations. Patron support and subsidies were rarer in Great Britain.

- Rest of Europe
  - **Russia.** The only Russian academy in my analysis is the one in Saint Petersburg. It was created under the direct patronage of Tsar Peter the Great, who fully financed and controlled it by retaining the right to approve new memberships. This academy was devoted to advancing Russia through the study of sciences and mathematics, along with history and humanities. The educational aspect was much more significant for this academy than for those in the rest of Europe.
  - **North Europe.** In the Netherlands, Sweden, Ireland, Belgium, Denmark, and Norway, official recognition was much more widespread than in the rest of Europe, with only a few informal or private academies. Nevertheless, the timing of granting recognition could vary from almost no wait to 19 years. This was the case for the Royal Dublin Society, which originated from a previous society in 1731 and obtained its royal charter only on April 2, 1750 (Berry, 1915). The topics of research were very similar to the rest of Europe, ranging from natural science to practical applications to improve local society.
  - **East Europe.**[51] In the Czech Republic both of the academies, in Prague and Olomouc, received official recognition, while in Switzerland the academies were mostly private societies (Kostlán, 1996; Zacek, 1968). Olomouc stood out for its open-minded atmosphere, where Catholics and Protestants collaborated and helped each other (Kostlán, 1996). The topics of study focused on natural history and improving the efficiency of agriculture, which was especially important for Switzerland (Rübel, 1947).

[51] McClellan (1985) does not include any academy in Poland; however, the "Warsaw Society of Friends of Science" (*Towarzystwo Przyjaciół Nauk*) was established in 1800 and remained active until 1832. It is the earliest scientific academy recorded in Poland. Nevertheless, I will not include it in the analysis, as data collection has only recently begun.

– **South Europe.** Both Spain and Portugal had one scientific academy each,[52] and both were officially recognized a year after their creation. Patronage was an important source of finance for these societies as well (Teixeira Rebelo da Silva, 2015). The topics were similar, focusing on the advancement of the local region.

2.A.2 Academies' establishments

- **Acad Agen.** The *"Société des Sciences, arts et belles lettres*" was created in Agen (FRA) on January 1, 1776 and officially recogniez in 1788. The main goal was to provide a forum for intellectual discussion. It was focused on advancing knowledge in various fields, including arts, sciences, and belles-lettres (Lauzun, 1900).

- **Acad Amiens.** The *"Académie des Sciences, belles lettres et arts"* was founded in Amiens (FRA) in February 1746 and officially recognized in 1750. It was initiated by Chauvelin German Louis (lawyer), Gresset Jean-Baptiste (poet), and d'Albert d'Ailly Michel Ferdinand (governor and scientist). It aimed to cultivate the spirit and shape taste through perfecting language, art, and knowledge. It had a structured governance with a hierarchy and specific tasks for its members (Académie des sciences, des lettres et des arts d'Amiens, 1901).

- **Acad Angers.** The *"Académie des sciences, belles lettres et arts d'Angers*" was founded in Angers (FRA) on March 31, 1684, and received official recognition through patent letters from Luis XIV in June 1685. It was established under the proposition of the mayor of Angers, Jacques Charlot (Bois, 2021).

- **Acad Arras.** The *"Académie Royale de belles lettres*" was founded in Arras (FRA) on May 22, 1737, but received official recognition with patent letters in 1773. It was founded by the writer Pierre Antoine de La Place, the military engineer Victor-Hyacinthe d'Artus, and the counsellor Galhaut de Lassus. The academy sought to advance knowledge in literature and the arts. It had a structured governance with different categories of members and a protector. Maximilien de Robespierre joined the Academy of Arras in

[52] I refer to the list in McClellan (1985), which includes only one Spanish and one Portuguese academy. However, we recently found two other scientific academies in Madrid. The oldest is *Real Academia de Matemáticas de Madrid* created in 1582 and hence out of the academy movement I study in this paper. However, it may be argued that it contributed to its origins. The other is the Royal Spanish Academy, founded in Madrid in 1713 and dedicated to the study of all sciences (Real Academia de Ciencias Exactas, Físicas y Naturales de España, 2024). Despite this, I cannot include these academies in my analysis, given we only started the search for reliable sources and data.

1783, highly increasing the academy reputation (Académie des Sciences, Lettres et Arts d'Arras, 2024).

- **Acad Arrezo.** The *"Accademia Aretina"* was established in Arezzo (ITA) in 1787. It was founded by a group of 22 scholars to revitalize the intellectual atmosphere that was lacking scientific and literary discussions following the closure of two previous academies (Maylender, 1930, Vol 1). It is still active today.

- **Acad Auxerre.** The *"Académie des Sciences, arts et belles lettres"* was founded in Auxerre (FRA) in April 1749 with the permission of the King and support from M. de Caylus, the bishop of Auxerre. The academy aimed to advance knowledge in various fields, including ecclesiastical, civil, and natural history, physics, and agriculture. Arts and literature were also included in their pursuits. It had a director and a perpetual secretary, similar to the Paris Academy (des Barres, 1851).

- **Acad Barcelona.** The *"Reial Acadèmia de Ciències i Arts de Barcelona"* was founded in Barcelona (ESP) on January 18, 1764, and officially recognized on December 17, 1765. It was established by 15 founders, led by Francesc Subiràs i Barra, the first director. The academy sought to spread scientific and technical knowledge to the city. It is still active today (Reial Acadèmia de Ciències i Arts de Barcelona, 2024).

- **Acad Berlin.** The *"Gesellschaft Naturforschender Freunde zu Berlin (GNF)"* was founded in Berlin (DEU) on July 9, 1773. It was established by the doctor and natural scientist Friedrich Heinrich Wilhelm Martini. The academy aimed to recruit and train young scientists, popularize science, and enhance the experimental study of natural history. It has a structured governance, including ordinary, honorary, and extraordinary members, and is still active today. The Prussian State was financing its activities (Böhme-Kaßler, 2005).

- **Acad Besançon.** The *"Académie des Sciences, belles lettres et arts"* was founded in Besancon (FRA) in 1748, and in 1752 officially recognized with patent letters from Luis XV. It was established by Pourroy de Quinsonas (president of the Franche-Comte Parliament), the duc de Tallard (governor of Comte du Bourgogne), and Moreau de Beaumont (intendant of Franche-Comte). The academy sought to create a lasting and organized forum for advancing knowledge in the sciences and arts. It had a structured governance with a protector and 40 titular members (Defrasne et al., 2002).

- **Acad Beziers.** The *"Académie des Sciences et belles lettres"* was founded in Beziers (FRA) on August 19, 1723, and became a Royal Academy in 1766 with the receipt of patent letters. It was established by the lawyer Antoine Portalon, the physicist Dortous de Mairan, and the doctor Jean Bouillet (Académie de Béziers, 2024).

- **Acad Birmingham.** The *"Lunar Society of Birmingham*" was founded in Birmingham (GBR) in 1766, following the "Lunar Circle" that formed in 1765. It primarily focused on science, both pure and applied, particularly as it related to industrial problems. Its members were mainly "provincial manufacturers and professional men."(Schofield, 1963, p.3) The academy did not have a structured governance. It was known for its monthly meetings held near the full moon (Schofield, 1963).

- **Acad Bologna.** The *"Istituto delle Scienze di Bologna*" was founded in Bologna (ITA) in 1714, although it had informal roots dating back to 1711 and a predecessor, the Inqueti Academy, established in 1690. The academy was founded by Count Luigi Ferdinando Marsili and Eustachio Manfredi, with papal patronage. It was established to foster reforms within the University. The academy's focus was on experimental sciences, medicine, physics, chemistry, and mathematics. It was the first Italian academy to have academicians employed and paid by public funds (Ercolani, 1881).

- **Acad Traccia.** The *"Accademia della Traccia*" was founded in Bologna (ITA) in 1666. It was established by Abate Carlo Sampieri, following the influence of Geminiano Montanari, a professor at the University of Bologna and a corresponding member of the Cimento Academy. The academy was created as an imitation of the Cimento Academy and focused on experimental physics (Maylender, 1930, Vol 5).

- **Acad Bordeaux.** The *"Académie royale des sciences, belles-lettres et arts*" was founded in Bordeaux (FRA) in 1712, through letters patent issued on September 5th. The academy's aim was to advance knowledge across a spectrum of disciplines, including belles-lettres, sciences, and arts. Natural history gained prominence with the establishment of the Société d'Histoire Naturelle in 1796. It had a structured governance with ordinary members, associate members, directors, secretaries, and a treasurer (Courteault, 1912).

- **Acad Bourg-en-Bresse.** The *"Académie des Sciences, belles lettres et arts*" in Bourg-en-Bresse (FRA) was initially established in 1755 and then reconstructed in 1783. It was founded by a group of notables to foster intellectual pursuits and the exchange of knowledge. The society focused on science, agriculture, letters, and social issues. It had a structured governance, with a Director, Vice-Director, and a perpetual Secretary (Allombert, 1899).

- **Acad Bratislava.** The *Pressburgischen Gesellschaft der Freunde der Wissenschaften* was the first Hungarian society. It was created in 1752 by a group of protestant scholars, guided by Karl Gottlieb Windisch, and it was active only until 1763 (Réka & Tüskés, 2017).

- **Acad Brest.** The *"Académie Royale de Marine*" was founded in Brest (FRA) in 1750 and officially recognized on July 30, 1752. It received royal status in 1769. It was founded by Sébastien Bigot de Morogues, a naval officer and scholar. The academy sought to study everything related to the navy, including naval officer training, shipbuilding techniques, research in mathematics, physics, arts, and natural history, and the compilation of a "Dictionary of Marine." It had a structured governance with honorary academicians, free/associate academicians, correspondents, ordinary academicians, and adjunct academicians (Académie Royale de Marine, 2024).

- **Acad Bruxelles.** The *"Académie Royale et Imperiale des Sciences et belles lettres*" was founded in Bruxelles (BEL) in 1769 as an informal society and officially recognized as a society with patent letters from Maria Theresa in 1772. It was founded by Count Cobenzl, who was inspired by the advice of Professor Schoëfflin. The academy aimed to revive interest in literature in the Austrian Netherlands, which was seen as declining. It had a structured governance with honorary members and ordinary academicians (Hasquin, 2009).

- **Acad Caen 1.** The *"Académie de physique de Caen*" was founded in Caen (FRA) in 1652, but received patent letters in 1705. It was established by Moisant de Brieux, de Grentemesnil, de Prémont, Halley, Vicquemand, and Bochart. The academy's initial focus was on literature and philosophy, but shifted to scientific matters after the creation of the Royal Society and the Academy of Sciences. It had a structured governance, with a director, a secretary, and a permanent reader (de Pontville, 1997).

- **Acad Caen 2.** The *"Académie des arts et belles lettres*" was founded in Caen (FRA) in 1662. It was never officially recognized by the King but the ministry Colbert expressed the royal approval. It was established by Pierre-Daniel Huet, who was inspired by the mostly literary works of other academies. The academy focused on physical and mathematical sciences. It had a structured governance and a clear set of objectives for its research (de Pontville, 1997).

- **Acad Châlons-en-Champagne.** The *"Académie des Sciences, arts et belles lettres*" was founded in Châlons-en-Champagne (FRA) in 1750, officially recognized in 1753, and received patent letters in 1775. Its motto was "L'Utilité," emphasizing practical applications of knowledge. It sought to cultivate belle-lettres, arts, sciences, and research in natural history. It had a structured governance with honorary academicians, titular academicians, "Agrégés pour les Arts," and associate free members (Menu, 1869).

- **Acad Cherbourg.** The *"Société Académique*" was founded in Cherbourg (FRA) on January 14, 1755, and officially recognized in 1775. It was established by Jean-François Delaville and other 5 scholars to share knowledge and improve the reputation of the city. It was focused on the history of the local region and archeology, and naval matters also entered into its discussions. It had a structured governance similar to the Paris Academy(Académie De Cherbourg, 2024).
- **Acad Cimento.** The *"Accademia del Cimento*" was founded in Firenze (ITA) in 1651 as an informal society and officially established in 1657. It was founded by Grand Duke Ferdinando II and his brother Leopoldo, who advocated for the free application of the "New Science." The academy focused on experimental physics, meteorology, and astronomy. It was considered innovative in its methodology (Knowles Middleton, 1971).
- **Acad Florence 2.** The *"Accademia Botanica*" was founded in Firenze (ITA) in 1733 and officially recognized in 1739. It was established by Vincenzo Capponi as secretary of the earlier Botanic Society. It focused on scientific research and studies, with a particular interest in botany and the management of the botanic gardens in Florence (Maylender, 1930, Vol 1).
- **Acad Florence 3.** The *"Reale accademia dei Georgofili*" was founded in Firenze (ITA) in 1753. The creation of the academy was prompted by an essay by Abbot Ubaldo Montelatici, who also incorporated the previous Botanic Academy of Florence in 1783. It was established to promote research in agronomy, especially to address issues with famine and food shortages in Italy. It had a structured governance (Tabarrini, 1856).
- **Acad Clermont Ferrand.** The *"Académie des Sciences, arts et belles lettres*" was founded in Clermont-Ferrand (FRA) in 1747, officially recognized in 1750, and granted patent letters in 1780. The academy was established by Rossignol, Dufraisse de Vernines, and Queriau, with the aim of promoting science and society through research in natural history and literature. The academy had a structured governance (Mège, 1884).
- **Acad Copenhagen.** Founded in Copenhagen (DNK) on November 13, 1742, and granted royal status in 1743, the *"Det Konelige Danske Vidensakkernes Selskab*" was established by Johan Ludvig Holstein, Hans Gram, Erik Pontoppidan, and Henrik Henrichsen. The academy aimed to strengthen

the position of science in Denmark and promote interdisciplinary understanding. It had a structured governance with a president and a secretary (Lomholt, 1950).

- **Acad Cosenza.** The *"Accademia dei Pescatori Cratilidi"* was founded in Cosenza (ITA) in 1753, inaugurated in 1756, and officially approved in 1758. It was established by Gaetano Greco, who wanted to create a new academy following the decline of the previous "Cosentina" academy. The academy's name derived from the Crati river and its motto was "Grandia ab exiguo" (i.e., "from small to large") (Maylender, 1930, Vol 4).

- **Acad Dantzig.** The *"Danziger Naturforschenden Gesellschaft"* was founded in Danzig (POL) on January 2, 1743. It was established during an informal gathering, with Daniel Gralath proposing the idea. The academy aimed to advance the understanding of natural phenomena through empirical investigation. It had a structured governance with different types of members and a permanent location (Schumann, 1893).

- **Acad Derby.** The *"Derby Philosophical Society"* was founded in Derby (GBR) in February 1783. It's heavily implied that Erasmus Darwin was a driving force behind the society, he was also member of the Lunar Society of Birmingham. The academy aimed to promote knowledge and discussion of natural philosophy and provide access to scientific literature through its library (Sturges, 1978).

- **Acad Dijon.** The *"Académie des Sciences, Arts et Belles-Lettres de Dijon"* was founded in Dijon (FRA) in 1725, established through the will of Hector-Bernard Pouffier (dean of the Parliament of Burgundy) and officially recognized by the King in 1740. It primarily focused on scientific subjects like medicine, natural sciences, and applied sciences, but also included a quarter of its members working in the humanities. It had a structured governance and was supported by the regional State of Burgundy (Milsand, 1871).

- **Acad Dublin.** The *"Philosophical Society and Medica-Philosophical Society"* in Dublin (IRL) evolved from a previous academy founded in 1683, which is from when I consider it active. However, it became the RDS (Royal Dublin Society) on June 25, 1731. It received royal recognition on April 2, 1750. The Dublin Society focused on improving the economy and the lives of the Irish people by promoting husbandry, manufactures, and useful arts. It had a structured governance with ordinary, honorary, and life members, and received funding through member subscriptions and parliamentary grants (Berry, 1915).

- **Acad Irish.** The *"Royal Irish Academy*" was founded in Dublin (IRL) in 1785 and granted a royal charter in 1786. This academy, the first in Ireland to balance research in both sciences and humanities, aimed to promote and investigate the sciences, polite literature, and antiquities. It had a structured governance with scientific and literary members, plus a rotating president (Royal Irish Academy, 2024).

- **Acad Edinburgh.** The *"Royal Society of Edinburgh*" was officially founded in Edinburgh (GBR) in 1783, with its first meeting on June 23, 1783. It received a Royal Charter on March 29. Many members of the earlier Philosophical Society became members of the RSE. This earlier society was founded in 1731, which will be the date from when I consider the academy active. It aimed to advance learning and useful knowledge, focusing on natural philosophy and literature. It had a structured governance (Emerson, 1981).

- **Acad Erfurt.** The *"Academia electoralis moguntina scientiarum utilium*" was founded in Erfurt (DEU) on July 19, 1754. Its creation was supported by its patron, the Elector of Mainz, Johann Friedrich Carl. The Academy aimed to promote useful sciences, like including natural sciences, mathematics, law, history, and the arts (Kiefer, 2004).

- **Acad Gorlitz.** The *"Oberlausitzischen Gesellschaft der Wissenschaften*" was founded in Gorlitz (DEU) on April 21, 1779. It was established by Karl Gottlob Anton, who proposed the idea to Adolf Traugott von Gersdorf. The academy aimed to promote the study of natural science and history in Upper Lusatia and foster scientific research and scholarship (Oberlausitzische Gesellschaft der Wissenschaften, 2024).

- **Acad Goteborg.** The *"Kungl. Vetenskaps-och Vitterbets Samhallet*" was founded in Goteborg (SWE) in the 1770s and obtained the royal title from King Gustav III in 1778. It was established by Johan Rosen, a schoolmaster, and later by Martin Georg Wallenstrale and Carl Fredrik Scheffer. The society aimed to promote scientific exchange among different disciplines and to foster the study of sciences for the benefit of the local community.

- **Acad Göttingen.** The *"Akademie der Wissenschaften zu Göttingen*" was established in Göttingen (DEU) in 1752 as the *"Königliche Societät der Wissenschaften*" (Royal Society of Sciences). It was founded under the patronage of King George II of Great Britain and Elector of Hanover. The academy aimed to advance learning and knowledge (Krahnke, 2001).

- **Acad Grenoble.** The *"Académie Delphinale*" was founded in Grenoble (FRA) in 1772, received patent letters in 1780, and formally adopted its name in

March 1789. It was established by a group of enlightened and noble men who purchased books following the death of the bishop of Grenoble. The academy focused on enhancing humanities like history, letters, and arts, but also included sciences and technical matters (*Lettres Patentes*, 1790).

- **Acad Haarlem.** The *"Hollandsche Maatschappij der Wetenschappen*" was founded in Haarlem (NLD) in 1752. It was established by seven leading citizens of Haarlem with the aim of promoting science. The academy has a twofold structure, with social members (representing society's interest in science) and scientific members (a group of scientists). It is still active today (Hollandsche Maatschappij der Wetenschappen, 2024).
- **Acad Tweede.** The *"Teylers Tweede Genootschap*" was founded in Haarlem (NLD) in 1756 and officially opened in 1778. It was established based on the will of Pieter Teyler van der Hulst. The academy aimed to promote science and the arts through discussion and prize competitions.
- **Acad Bad-Homburg.** The *"Société patriotique de Hesse-Hamburg pour l'encouragement des connaissances et des moeurs*" was founded in Bad-Homburg (DEU) in 1775, with statutes adopted in 1777. The academy aimed to promote "knowledge and morals" (from the name of the academy) and therefore focused on intellectual and ethical development (1777).
- **Acad Investiganti.** The "*Accademia degli Investiganti*" was founded in Napoli (ITA) in 1650 by Cornelio Tommaso and di Capua Leonardo. It was inspired by the Lincei academy in Rome, and sought to study and investigate "things of nature." It primarily focused on natural philosophy before 1735 and on literary matters after that (Maylender, 1930, p.369, Vol3).
- **Acad Naples.** The *"Reale Accademia della Scienze e Belle-Lettere*" was founded in Napoli (ITA) in 1778 and officially established in 1780. It was established by King Ferdinando IV of Borbon to advance public education, progress, and human conviction. It had a structured governance with a president, vice-president, treasurer, fiscal lawyer, and secretary, and received financial support from a royal annuity (Maylender, 1930, Vol 5).
- **Acad Jena.** The *"Naturforschende Gesellschaft zu Jena*" was founded in Jena (DEU) in 1793 by August Johann Georg Karl Batsch. The academy aimed to support members in choosing a career through natural-historical studies and to contribute to their moral advancement (Böhme-Kaßler, 2005).
- **Acad La Rochelle.** The *"Académie Royale des Belles lettres*" was founded in La-Rochelle (FRA) in 1730 and officially recognized in 1744. It was founded by Jean-Jacques Franc de Pompignan, who was considered the

soul of the academy. The academy focused on the study of literature and eloquence, specifically poetry. It had a structured governance with a director and a permanent secretary (Flouret, 2009).

- **Acad Lausanne.** The *"Société des sciences physiques*" was founded in Lausanne (CHE) on March 10, 1783. It aimed to cultivate interest in natural history and to study all that concerns the sciences, arts, agriculture, industry, commerce, and the local patrimony 1789.

- **Acad Leipzig.** The *"Fürstlich Jablonowskische Gesellschaft*" was founded in Leipzig (DEU) in 1768. Further sources have been asked to the current academy.

- **Acad Leopoldina.** The *"Deutsche Akademie der Naturforscher Leopoldina*" was founded in Halle (DEU) on January 1, 1652, and officially recognized by the Emperor Leopold I in August 1677. It was established by four physicians: Bausch, Fehr, Metzger, and Wohlfahrth. The academy aimed to explore nature for the glory of God and the good of mankind. It had a structured governance and received special privileges from the Emperor Leopold I (Deutsche Akademie der Naturforscher Leopoldina, 2024).

- **Acad Halle.** The *"Gesellschaft der Naturforschenden Freunde*" was founded in Halle (DEU) in 1779 by some theology students with the support of Friedrich-Wilhelm von Leysser, who became the first president. The academy aimed to increase acceptance and interest in natural history among students (Böhme-Kaßler, 2005).

- **Acad Lisbon.** The *"Academia real das ciencias de Lisboa*" was founded in Lisboa (PRT) in 1779 and officially recognized by the King in 1780. It was established by the Duke of Lafões, who provided significant financial support. The academy aimed to promote scientific knowledge and cultural development within Portugal. It had a structured governance and was primarily funded through royal patronage and private donations (Teixeira Rebelo da Silva, 2015).

- **Acad Lund.** The *"Kungl Fysiografiska Sallskapet*" was founded in Lund (SWE) in 1772, and officially recognized by King Gustav III in 1788. It was established by Theologian Hesselen, doctor in Medicine Barfort, and Magistrat Retzius. The academy aimed to encourage a passion for science in youth and to associate those who shared this passion to produce useful findings for the general public. It was devoted to natural history and economics (Gertz, 1940).

- **Acad Lyon.** The *"Académie Royale des Sciences, belles-lettres et arts de Lyon*" was founded in Lyon (FRA) in 1700, and officially recognized with patent letters in 1724. It was established by Claude Brossette and other notable citizens, aiming to promote the advancement of science, art, and literature in Lyon and the region. It had a structured governance with a director and a vice-director (Académie Royale des Sciences, belles-lettres et arts de Lyon, 2024).

- **Acad Manchester.** The *"Literary and philosophical society"* was founded in Manchester (GBR) in 1781. The academy was established by Thomas Percival and a group of men who sought to improve the living standards of the city, especially for the working class. It aimed to improve the local society and bring it towards more unity and progress (1896).

- **Acad Mannheim 1.** The *"Academia Electoralis Scientiarum et Elegantiorum Literarum Theodoro-Palatina"* was founded in Mannheim (DEU) between October 15–20, 1763. The academy was established by Karl Theodor, the Elector Palatine of Bavaria, and effectively organized by the French historiographer Johann Daniel Schöpflin and Leopold Maximilian, Baron of Hohenhausen, the Prince's chamberlain. The academy aimed to promote both science and the humanities. It had a structured governance system with a president and a secretary appointed for life (Academiae Electoralis Scientiarum et Elegantiorum Literarum Theodoro-Palatina, 1766; Cassidy, 1985). Another relevant institution, the *"Societas Meteorologica Palatina"*, was also founded in Mannheim (DEU), on September 5, 1780. It was likewise initiated and supported by Elector Palatine Karl Theodor and strongly advocated by scholars Father Hemmer and Stefan von Stengel. This academy focused on meteorology, aiming to connect international meteorological stations equipped with similar instruments to allow for comparative measurements. It, too, had a structured governance system and received financial support from Karl Theodor (Cassidy, 1985). I consider these two academies as a unique one, especially given the presence of intermediary societies and institutions that bridged the two, paving the way for the latter meteorological society.

- **Acad Mantua.** The *"Accademia Virgiliana"* was founded in Mantova (ITA) in 1686. It took the name *"Royal Academy of Sciences, Lettres, and Arts"* in 1768. It was established by the co-regnant Maria Teresa and Giuseppe II, with the aim of continuing intellectual development in the Austrian Lombardy. It initially focused on theology and letters, but later expanded to include sciences useful to society. It had a structured governance with members and a patron (Maylender, 1930, Vol 5).

- **Acad Marseille.** The *"Académie des belles-lettres, sciences et arts"* was founded in Marseille (FRA) in August 1726 and officially recognized by King Luis XV with patent letters in 1766. The academy's primary goal was to promote French language and literature in the region. It had a structured governance (Académie des Sciences Lettres et Arts de Marseille, 2024).

- **Acad Messina.** The *"Accademia Peloritana dei Pericolanti*" was founded in Messina (ITA) in 1728. It was established by Paolo Aglioti and others, following the death of Pietro Guerriera who had initially pushed for a similar academy. The academy focused on Letters, Moral and Natural Philosophy but also on Mathematics, Geography,and Duel and Knights subjects. After 10 years of activity, it focused primarily on scientific matters. It had a structured governance (Accademia Peloritana dei Pericolanti, 2024).

- **Acad Metz.** The *"Société Royale des Sciences et Arts*" was founded in Metz (FRA) in April 1757 and received patent letters in July 1760. The Marshal-Duke Charles Louis Auguste Fouquet de Belle-Isle was its founder and protector. The academy aimed to advance sciences, letters, and arts to make them useful to the local society of Metz.

- **Acad Middelburg.** The *"Zeeuwsch Genootschap der Wetenschappen*" was founded in Vlissingen (NLD) in 1765 and officially founded in 1769. It was established to provide a local organization for scientific practice and to promote the ideas of the Enlightenment (Zeeuwsch Genootschap der Wetenschappen, 2024).

- **Acad Modena.** The *"Accademia ducale dei Dissonanti di Modena*" was founded in Modena (ITA) in 1680 and formally active in 1684. It was established by the citizens of Modena to ask for the reopening of the University and the creation of the Academy. The academy was initially active only in humanities and letters, but added a scientific section in 1790 (Accademia Nazionale di Scienze, Lettere e Arti di Modena, 2023).

- **Acad Rangoniana.** The *"Accademia Rangoniana*" was founded in Modena (ITA) in 1783. It was established by Gherardo Aldobrandino Rangone, who was already financing and hosting scientific experiments of Michele Rosa, who worked on blood transfusions among animals. The academy focused on scientific experiments, mechanics, and physics (Maylender, 1930, Vol 4).

- **Acad Montauban** The *"Académie des belles lettres*" was founded in Montauban (FRA) in 1730 and officially recognized in 1744. The soul of the academy was Jean-Jacques Franc de Pompignan. The academy focused on literary subjects, particularly poetry and letters (Forestié, 1888).

- **Acad Montpellier** The *"Société Royale des Sciences*" was founded in Montpellier (FRA) in 1706. The King wanted to reassure his domain into the Mediterranean coast during the Spanish Succession. It was initially focused on mathematics, anatomy, chemistry, botany, and physics. It played a role in compiling the Encyclopédie of Diderot and d'Alembert (Dulieu, 1975; Société Royale des Sciences, 2024).

- **Acad Munich** The *"Bayerische Akademie der Wissenschaften*" was founded in Munchen (DEU) on October 12, 1758 and officially recognized on June 25, 1759. It was established by Johann Georg von Lori and aimed to advance all useful sciences in Bavaria (Bayerische Akademie der Wissenschaften, 2024).

- **Acad Nancy** The *"Société des Sciences et belles lettres - Académie Stanislas*" was founded in Nancy (FRA) on December 28, 1750, and received patent letters on December 27, 1751. It was founded by Stanislas Leszczynski, the king of Poland and duke of Lorraine and Bar. It aimed to enhance the study of sciences and literature and culture. It created a public library too (Stanislas, 2024).

- **Acad Nimes** The *"Academie Royale de Nimes*" was founded in Nimes (FRA) on March 28, 1682, and received patent letters in August 1682 from Luis XIV. It was established by Jules de Fayn, and aimed to enhance the local patrimony by studying antiquities and the local language (Nicolas, 1854).

- **Acad Nuremberg** The *"Cosmographical Society*" was founded in Nurnberg (DEU) in 1747.

- **Acad Olmouc.** The *"Societas Eruditorum Incognitorum*" was founded in Olomouc (CZE) in 1747 by Josef Petrash, who had traveled the world as a soldier and poet. The academy aimed to free higher education from the influence of Jesuits. It sought to cultivate the fine sciences and liberal arts (Kostlán, 1996).

- **Acad Orleans.** The *"Académie Royale des Sciences, arts et belles lettres*" was founded in Orleans (FRA) on April 23, 1781, and received patent letters on March 20, 1784. The academy was established by a group of 10 scholars. It aimed to promote physics and natural sciences (Nicolas, 1908).

- **Acad Oxford.** The *"Oxford Philosophical Society*" was founded in Oxford (GBR) in 1645 as an informal society and formally established in 1651 by John Wilkins and other natural philosophers. It was inspired by the London group of natural philosophers, and the remnants of William Harvey's circle at Oxford. The academy focused on magnetic experiments, dissections, antiquities, astronomy, and geometry (Applebaum, 2000; Gunther, 1925).

- **Acad Padua.** The *"Accademia dei Ricovrati/Accademia di Scienze, lettere ed Arti*" was founded in Padova (ITA) in 1599, which is still considered a Reinassance Academy (McClellan, 1985). It became the "Accademia di Scienze, lettere ed Arti" in 1779, when the Venetian Senate ordered its fusion with the *Accademia di Arte Agraria*. It was founded by Federico Cornaro, and Galileo was a founding member of the earlier Ricovrati Academy. The academy aimed to promote the study of humanities and science via the experimental approach (Maggiolo, 1983). The academy enter into my analysis only from 1779.

- **Acad Palermo.** The *"Accademia Palermitana*" was founded in Palermo (ITA) in 1718, though it only received recognition in 1752. It was established by Pietro Filangieri and other enlightened men. The academy aimed to tell Sicily's story and advance letters and sciences (Maylender, 1930, Vol 1).

- **Acad Palma.** The *"Accademia Boreliana*" was founded in Palmi (ITA) in 1673 by Gio. Alfonso Borelli. It focused on physics and natural history, especially on the respiration moto Maylender (1930, Vol 1).

- **Acad Pau.** The *"Académie Royale des Sciences et beaux arts*" was founded in Pau (FRA) in 1718.

- **Acad Prussia.** The *"Königlich-Preußische Akademie der Wissenschaften*" was founded in Berlin (DEU) on July 11, 1700, and immediately officially recognized. It was established by Gottfried Wilhelm von Leibniz, with sponsorship from the noble Hohenzollern family. The academy aimed to advance both humanities and natural sciences (de la Croix, Eisfeld, & Ganterer, 2021; Königlich-Preußische Akademie der Wissenschaften, 2024).

- **Acad Prague.** The *"Regia Societas Scientiarum Bohemica*" was founded in Praha (CZE) in 1769 and officially recognized in 1790. The academy was established by count Frantisek Josef Kinsky and Ignac Born. It aimed to diffuse the experimental approach and critical thinking but also Bohemian History (Zacek, 1968).

- **Acad Reggio d'Emilia.** The *"Accademia degli Ipocandriaci*" was founded in Reggio-Emilia (ITA) in 1746. It was established by Achille Crispi, the captain of the Duke Francesco III. The academy had a structured governance (Maylender, 1930, Vol 3).

- **Acad Roma.** The *"Accademia di Fisico-Mathematica*" was founded in Roma (ITA) on July 6, 1677. It was established by Giovanni Giustino Ciampini, who provided the academy with tools and machines for scientific experiments. The academy focused on natural sciences and experiments, including anatomy, physics, mathematics, and mechanics (Maylender, 1930, Vol 3).

- **Acad Rotterdam.** The *"Battafsch Genootschap der Proefonderwindelijke Wijsbegeerte*" was founded in RotterdamNLD on May 14, 1769 (Lieburg, 1985).

- **Acad Rouen.** The *"Académie Royale des Sciences, belles lettres et arts*" started informally in Rouen (FRA) in 1736 and formally with patent letters from Luis XV on June 17, 1744. It was established by Fontanelle and Le Cornier de Cideville, and focused on botany (Gosseaume, 1985).

- **Acad Rovereto.** The *"Imperiale Regia Accademia degli Agiati*" was founded in Rovereto (ITA) in 1750, officially recognized in 1753. It was established by Giuseppe Valeriano Vannetti and other four important local scholars. The academy was initially focused on letters, history, and science, but later expanded to include agricultural research (Accademia Roveretana degli Agiati di Scienze, Lettere ed Arti, 2024).

- **Acad Paris.** The *"Académie Royale des Sciences*" was founded in Paris (FRA) in the spring of 1666. The academy was established by Minister Colbert under Luis XIV, who fully funded its creation and operations. The academy was a symbol of royal patronage. Its focus was on natural philosophy, mathematics, and the application of the laws of nature to practical reforms (Académie Royale des sciences, 2024).

- **Acad Siena.** The *"Reale Accademia della scienze di Siena*" was founded in Siena (ITA) in 1690. It was established by Pirro Maria Garieli, a professor at the University of Siena. The academy focused on natural science, philosophy, medicine, and poetry Maylender (1930, Vol 3).

- **Acad Spalding.** The *Philosophical Society* was founded informally by Maurice Johnson in 1710 to foster the study of archaeology and antiquarianism. Two years later the organization of the society became more formal with structured meetings and transcription of the minutes (Spalding Gentlemen's Society, 2025).

- **Acad Stockholm.** The *"Kungliga Vetenskapsakademien*" was founded in Stockholm (SWE) on June 2, 1739. It was modelled after the Royal Society of London and the Académie Royale des Sciences in Paris. The academy was created as an independent, non-governmental scientific society. It was primarily focused on natural sciences and mathematics (Kungliga Vetenskapsakademien, 2024).

- **Acad St Petersburg.** The *"Academia Scientiarum Imperialis Petropolitanae*" was founded in Saint-Petersburg (RUS) in 1724. It was established by Peter the Great, who was inspired by academies in Europe. The academy aimed to bring the Russian Empire into the modern era. It was initially focused on mathematics, physical sciences, and humanities, and included training in scientific subjects (de la Croix & Doraghi, 2021; Gordin, 2000).
- **Acad Toulouse.** The *"Académie Royale des Sciences, inscriptions et belles lettres*" was founded in Toulouse (FRA) in 1640-1645/1665-1685 as an academic conference and officially recognized in 1746. It was established by Sage Antoine, Carrière, and Gouazé Pierre. The academy aimed to advance sciences, inscriptions, and belles-lettres. It had a structured governance (Taillefer, 1984).
- **Acad Trondheim.** The *"Det Kongelige Norske Vienskabers Selskab*" was founded in Trondheim (NOR) in 1760 and received royal recognition in 1767. It was established by Bishop Johan Ernst Gunnerus, rector Gerhard Schoning, and councilor Peter Friderich Suhm to create an institutional space for enhancing and spreading the New Science (Schmidt, 1960).
- **Acad Turin.** The *"Accademia delle scienze di Torino*" was founded in Torino (ITA) in 1757 and officially recognized in 1783. It was founded by Joseph-Louis Lagrange, Giuseppe Francesco Cigna, and Giuseppe Angelo Saluzzo. The academy aimed to advance scientific research that could not find enough space within the university of the city (Accademia delle Scienze di Torino, 2023).
- **Acad Filopatria.** The *"Accademia Filopatria*" was founded in Torino (ITA) on July 2, 1782. It was established by a group of enlightened men in the city of Turin. The academy focused on antiquities and the history of the homeland, including letters, poetry, and moral values but also on public economics, and sciences (Campori, 1887).
- **Acad Uppsala.** The *"Societatis Regiae Scientiarum Upsaliensis*" was founded in Uppsala (SWE) in 1710. It was reorganized in 1719, and received royal recognition on November 11, 1728. It was founded by the librarian Eric Benzelius. The academy initially focused on scientific discussions and later established a scientific journal (Karlberg, 1977).
- **Acad Uppsala.** The *"Cosmographiska sällskapet*" was founded in Uppsala (SWE) in 1758 by Anders Akerman and other enlightened men. The academy focused on cosmography, constructing globes for the earth and the sky.

- **Acad Utrecht.** The *"Provinciaal Utrechtsh genootschap van Kunsten en Wetenschappen*" was founded in Utrecht (NLD) in 1773 and officially founded in 1778. It was established by Mr. J. van Haeften and L. Praalder. The academy aimed to preserve local heritage, modern art, and publications, as well as to develop and improve science (Singels, 1923).
- **Acad Valence.** The *"Société Académique et Patriotique*" was founded in Valence (FRA) in 1784, receiving King's Letters Patent in December 1786. The academy aimed to advance sciences, arts, and belles-lettres. It had a structured governance and it organized 3 prizes every year (de Colonjon, 1866).
- **Acad Venice.** The *"Accademia dei Planomaci*" was founded in Venezia (ITA) circa 1740. It was established by the abate D. Meodoro Rossi. The academy published the *"Novelle Letterarie,*" a journal of reviews and critiques of new works. It had a structured governance with a protector (Maylender, 1930, Vol 4).
- **Acad Verona.** The *"Societa Italiana delle Scienze*" was founded in Verona (ITA) in 1766 and officially established in 1782. It was founded by Antonio Mario Lorgna. The academy focused on scientific matters and published the periodical *"Memorie accademiche*" (Maylender, 1930, Vol 1).
- **Acad Zurich.** The *"Naturforschende Gesellschaft*" was founded in Zurich (CHE) in 1745 and formally established in 1746. It was established by Johanes Gessner. The academy aimed to provide a space for students and personalities who studied abroad to return home and share their knowledge. It had a structured governance and relied heavily on member contributions (Rübel, 1947).
- **Royal Society.** The *"Royal Society of London*" was founded in London (GBR) in 1660 and officially established in 1662. It was established by John Wilkins and other polymaths. The academy focused on natural philosophy and experiments, including trade, manufacture, and crafts, as well as scientific experiments. It had a structured governance with a president, a treasurer, and two secretaries (The Royal Society, 2024).
- **Acad Botanical.** The *"Botanical Society*" was founded in London (GBR) in 1721. It was established by Johann Jakob Dillen and John Martyn to increase knowledge of and spread interest in minerals, plants, and animals.
- **Acad Linnaus.** The *"Linnean Society of London*" was founded in London (GBR) in 1788 by James Edward Smith, Samuel Goodenough, and Thomas Marsham. The academy was named after Carl Linnaeus, who is considered

the father of taxonomy. The academy was devoted to natural history, focusing on the evolution theory and biological taxonomy.

2.B Descriptive Statistics

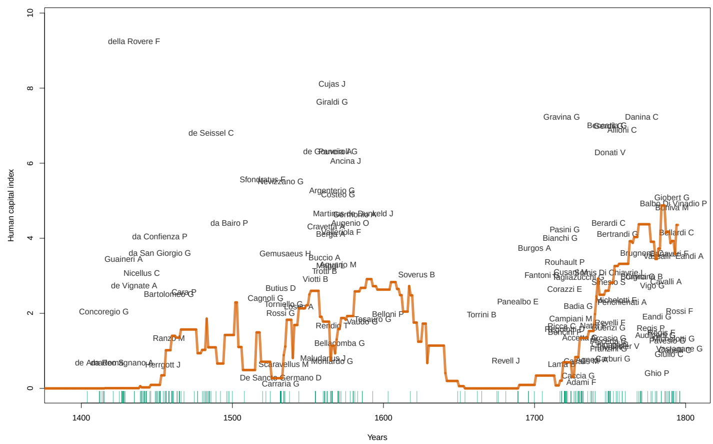


FIGURE 2.8: Notable scholars at the University of Turin.

The figure displays ordinary members of the University of Turin with a non-negative individual quality index. Vertical green lines represent the distribution of scholars over time, including those with a quality index of zero. The orange line shows the evolution of the university's aggregate quality. Source: Zanardello (2022).

TABLE 2.3: Summary Statistics: Founders vs. Non-Founders of Academies.

	(1) **Founders**	(2) **Not Founders**	(3) **t-test**
Obs.	413	16860	
	μ	μ	p-value
Quality	2.67	2.41	0.426
Age at death	68	67	0.189
Age at Appointment	36.7	37.3	0.611
Years Active	19.5	16	0.338
Distance Birthplace-Academy	220	344	0.003
Distance Academy-Death place	317	421	0.260
Distance Birthplace-Death place	248	368	0.074
Year FE*			YES

Note: Column (1) reports summary statistics for scholars who founded an academy; Column (2) shows statistics for those who did not. Column (3) reports p-values from t-tests comparing the means between the two groups, controlling for year fixed effects.
*Year fixed effects refer to the scholar's initial year of activity.

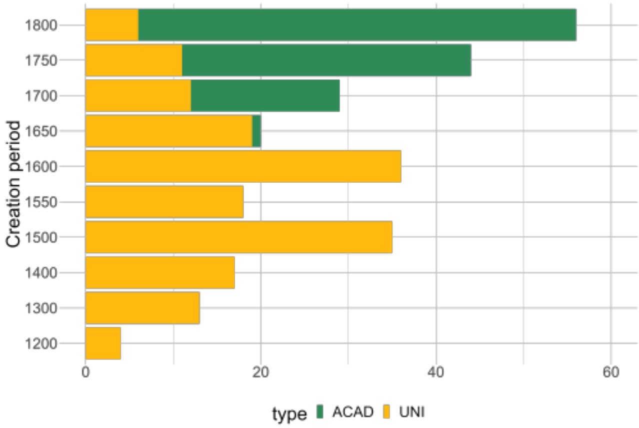


FIGURE 2.9: Creation Timeline of Academies and Universities.

Note: The figure shows the cumulative number of institutions established up to the year indicated on the y-axis. Universities (yellow) exhibit a more heterogeneous timeline, with establishments beginning in the 11th century. In contrast, academies (green) began to proliferate in the second half of the 17th century.

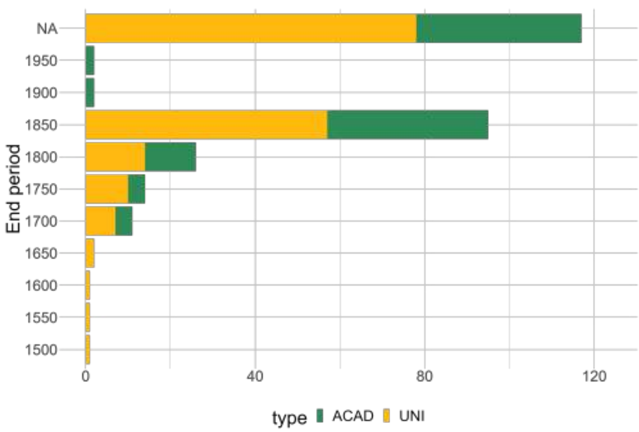


FIGURE 2.10: Closure Timeline of Universities and Academies.

Note: The figure shows the number of institutions that closed in the 50 years preceding the year indicated on the y-axis. Universities are shown in yellow; academies in green. *NA* refers to institutions that never closed. Most institutions remained open before 1800; a wave of closures occurred between 1800 and 1850, although the majority are still active today.

TABLE 2.4: Summary Statistics for Academies.

	Mean	**Med**	**Min**	**Max**	**Obs**
Total activity years	137	81	2*	372**	101
Creation Date	1740	1750	1650	1793	101
End Date	1877	1794	1667	2024	101
Total size	137	57	1†	1622‡	101
% SCIENCE	53.2%	53.0%	0.0%	100.0%	101
% LITERARY	43.1%	43.1%	0.0%	100.0%	101
% UNKNOWN	2.4%	0.0%	0.0%	28.7%	101
Size in 1650–1700	238	49	2	1622‡	16
% SCIENCE	58%	56.4%	19.6%	100.0%	16
% LITERARY	39.3%	41.8%	0.0%	73.9%	16
% UNKNOWN	2.6%	0.0%	0.0%	14.0%	16
Size in 1700–1750	142	70	4	873⊕	33
% SCIENCE	49.1%	47.4%	0.0%	100.0%	33
% LITERARY	49.5%	51.7%	0.0%	100.0%	33
% UNKNOWN	1.4%	0.0%	0.0%	12.9%	33
Size in 1750–1800	103	52	1†	480•	52
% SCIENCE	54.4%	54.3%	0.0%	100.0%	52
% LITERARY	42.6%	42.1%	0.0%	100.0%	52
% UNKNOWN	3%	0.0%	0.0%	28.7%	52

Note: This table presents summary statistics for the academies in my sample. 'Med' indicated the median, 'Min' the minimum, 'Max' the maximum, and 'Obs' the number of observations considered. * Refer to *Accademia della Traccia* in Bologna (ITA, 1665).
** Refer to *Leopoldina Academy* in Halle (DEU, 1652).
† Refer to *Naturforschende Gesellschaft* in Jena (DEU, 1793).
‡ Refer to the *Royal Society* in London (UK, 1660).
⊕ Refer to the *Prussian Academy* in Berlin (DEU, 1700).
• Refer to the *Erfurt Academy* in Germany (DEU, 1754).

TABLE 2.5: Summary Statistics for Universities.

	Mean	Med	Min	Max	Obs
Total activity years	366	356	3*	936**	171
Creation Date	1517	1548	1088	1781	171
End Date	1883	1811	1460	2024	171
Total size	155	70	0†	1958‡	171
% SCIENCE	20.7%	19.1%	0.0%	100.0%	171
% LITERARY	74.7%	76.9%	0.0%	100.0%	171
% UNKNOWN	2.8%	0.0%	0.0%	31.3%	171
Size in 1650–1700	91	50	12	246°	12
% SCIENCE	20.8%	19.2%	0.0%	40.0%	12
% LITERARY	74.3%	75.5%	60.0%	84.4%	12
% UNKNOWN	4.9%	1.7%	0.0%	16.7%	12
Size in 1700–1750	65	26	1	354⊕	11
% SCIENCE	28.8%	19.1%	0.0%	67.8%	11
% LITERARY	69.6%	81.0%	19.4%	100.0%	11
% UNKNOWN	1.6%	0.0%	0.0%	12.9%	11
Size in 1750–1800	27	20	1	71•	6
% SCIENCE	34.2%	33.6%	0.0%	80.0%	6
% LITERARY	65.5%	65.7%	20.0%	100.0%	6
% UNKNOWN	0.2%	0.0%	0.0%	1.4%	6

Note: This table presents summary statistics for the universities in my sample. 'Med' indicated the median, 'Min' the minimum, 'Max' the maximum, and 'Obs' the number of observations considered. * Refer to *Corte University* in France (1765).
** Refer to the *University of Bologna* in Italy (1088).
† Refer to the universities for which we did not find any member yet in Burgo-de-Osma (ESP, 1555), Genova (ITA, 1471), and Palma (ESP, 1483).
‡ Refer to the *University of Cambridge* in UK (1209).
° Refer to the *University of Lund* in Sweden (SWE, 1666).
⊕ Refer to the *University of Göttingen* in Germany (1734).
• Refer to the *University of Moscow* in Russia (1755).

(A) 1500 - 1550

(B) 1550-1600

(C) 1600 - 1650

(D) 1650 - 1700

(E) 1700 - 1750

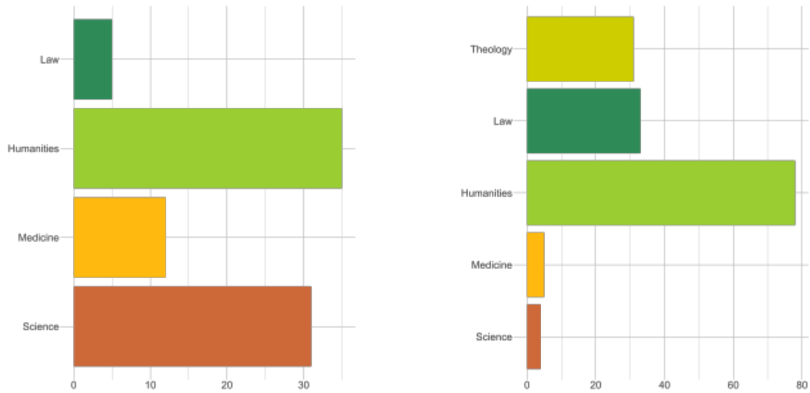
(F) 1750 - 1800

(G) 1800 - 1850

(H) 1850 - 1900

FIGURE 2.11: Locations of Academies and Universities (1500 - 1900 CE).

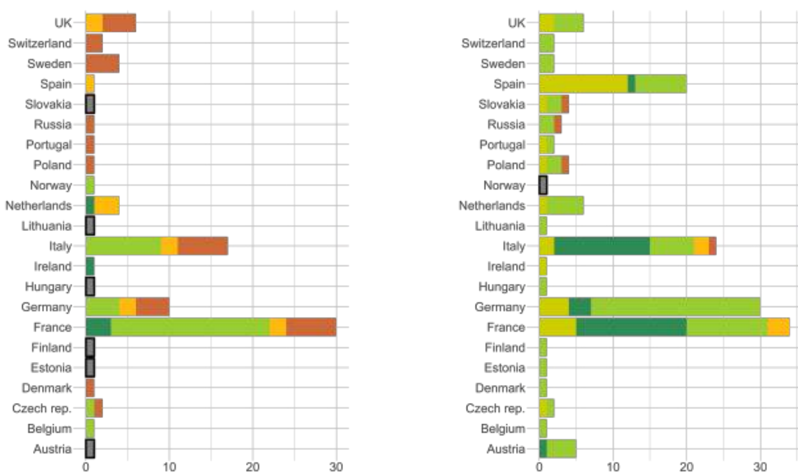
Note: Yellow circles represent universities, while green triangles represent academies. In cities where both institutions were established, the shapes overlap to indicate interaction. Country borders reflect those in the year 2000.



(A) Academies (B) Universities

FIGURE 2.12: Institutions by Main Field of Study.

Note: Number of institutions by their main field of study, defined as the field in which the majority of the institution's members were active.



(A) Academies (B) Universities

FIGURE 2.13: Institutions by Major Field of Study per Country.

Note: Number of institutions by major field of study, shown by country. The major field of study refers to the most commonly studied field within each institution, without applying a minimum threshold. Science is shown in red, Medicine in gold, Humanities in light green, Law in dark green, and Theology in lime. Countries without any identified institutions are shown in grey and outlined with a solid black border.

TABLE 2.6: Types of Academies by Country.

	(1) ACAD	(2) Scientific ACAD*	(3) Literary ACAD**	(4) Long-lasting ACAD	(5) Big ACAD
Europe	81	43	38	68	62
France	29	11	18	23	22
Italy	17	8	9	15	12
Germany	10	6	4	9	7
UK	6	6	0	5	5
Belgium	1	1	0	0	1
Czech Republic	2	1	1	1	1
Denmark	1	0	1	1	1
Ireland	1	0	1	1	1
Netherlands	4	2	2	4	3
Norway	1	0	1	1	1
Poland	1	1	0	1	1
Portugal	1	0	1	1	1
Spain	1	1	0	1	1
Sweden	4	4	0	4	3
Switzerland	2	2	0	1	2

Note: This table shows the number of the academies in my sample by type and by country. Columns (2) and (3) add up to Column (1). Columns (4) and (5) are independent of each other.
* An academy (ACAD) is classified as scientific if at least 50% of its members study science, applied science, or medicine.
** An academy is classified as literary if at least 50% of its members study theology, law, humanities, or social sciences.

TABLE 2.7: Pre-treatment Summary Statistics.

$EVENT$	Obs	Mean	St. Dev.	Min	Max
Outcome variable: $\Delta lnpop$ 1500-1900, 50-years interval					
Academy	81	0.204	0.302	−0.318	1.139
University	83	0.125	0.242	−0.434	1.099
Scientific Academy	43	0.205	0.312	−0.318	1.139
Literary Academy	38	0.203	0.297	−0.236	1.099
Long Academy	68	0.220	0.294	−0.318	1.099
Big Academy	62	0.224	0.320	−0.318	1.139
Outcome variable: $AvgQ$ 1500-1900, 50 years interval					
Academy	40	2.659	2.362	0	7.060
Scientific Academy	21	3.270	2.345	0	7.060
Literary Academy	19	1.984	2.249	0	6.768
Long Academy	35	2.814	2.314	0	7.060
Big Academy	32	2.549	2.271	0	7.060

Note: There are 149 universities in total, of which 66 opened before 1500 — the start of the sample period. Therefore, the statistics are computed using only 83 observations.

2.C Additional results

2.C.1 OLS results

In this section, I present results from OLS regressions examining the relationship between the natural logarithm of population and the presence of an academy, while controlling for the presence of a university in a European city during the period of analysis (i.e., 1500–1900). Table 2.8 reports estimates for the full sample of 2,096 cities, following Buringh (2021), while Table 2.10 focuses on a subset of 633 large cities as defined by Bosker, Buringh, and Van Zanden (2013).[53] In my preferred specification (Column 4), I include both city fixed effects and country-by-time fixed effects. This helps account for unobserved, time-invariant city characteristics and broader national trends. While no time-varying controls are available for the full sample, I argue that the combination of city and country-by-time fixed effects mitigates omitted variable bias. In Appendix 2.C.2, I support this claim using the smaller sample of large cities for which time-varying controls are available, showing that the results are similar with and without these controls (see Table 2.10).

Table 2.8 shows that the presence of an academy is positively and significantly associated with higher population growth. The magnitude of the academy coefficient (ACAD) is consistently larger than that of universities, suggesting a strong correlation between academies and urban growth. However, the interaction between academies and universities (ACAD × UNI) is not statistically significant in Columns 2 and 4. Interestingly, this coefficient is negative, suggesting that establishing an academy in a city that already has a university may not further enhance local economic conditions—and might even have a slight crowding-out effect. Nevertheless, the total correlation of having both institutions remains positive and significant. Results for the large-city sample in Table 2.10 mirror these patterns. Specifications that omit city fixed effects but include country and time fixed effects separately (e.g., Column 3 in Table 2.8 and Column 2 in Table 2.10) produce stronger coefficients, but they fail to account for city-specific factors, as reflected in lower R-squared values.

Table 2.9 presents a first heterogeneity analysis by field of study, as introduced in Section 2.3. The results suggest that the positive effects identified in Table 2.8 are primarily driven by scientific academies. Only the coefficients for academies focused on science, applied science, and medicine (Columns 1 and 3) are positive and highly significant. In contrast, literary academies—those emphasizing humanities, theology, law, and the social sciences (Columns 2 and 4)—do not yield significant results.

[53] Bosker, Buringh, and Van Zanden (2013) define large cities as those exceeding 10,000 inhabitants at least once between 800 and 1800. My full sample includes cities exceeding 5,000 inhabitants at least once between 700 and 2000.

TABLE 2.8: OLS Estimates: Effect of Academies and Universities on City Population.

	Dependent Variable: ln pop in 1500-1900			
	Obs.: **ALL cities** as in Buringh (2021)			
	(1)	(2)	(3)	(4)
ACAD	0.308**	0.316**	1.757***	0.271**
	(0.118)	(0.130)	(0.226)	(0.113)
UNI		0.142**	1.089***	0.082
		(0.056)	(0.077)	(0.056)
ACADxUNI		−0.055	−1.005***	−0.008
		(0.225)	(0.323)	0.205
Cities	2096	2096	2096	2096
R^2	0.806	0.806	0.442	0.845
city FE	YES	YES	NO	YES
time FE	YES	YES	YES	NO
country FE	NO	NO	YES	NO
countryXtime FE	NO	NO	NO	YES

Note: *p<0.1; **p<0.05; ***p<0.01.
Standard errors are clustered at the city level (in parentheses).
ACAD indicates the presence of only an academy in a given city and time period; **UNI** indicates the presence of only a university; **ACAD**×**UNI** denotes the presence of both an academy and a university in the same city and time period.
The dependent variable is the natural logarithm of city population, measured every 50 years from 1500 to 1900, following the dataset in Buringh (2021).
Each column reports estimates from a different specification. The models include the indicated combinations of fixed effects (city, time, country, or country-by-time).

The interaction terms between the presence of a university and either type of academy mirror those in Table 2.8: negative in sign and statistically insignificant, suggesting that the combination of both institutions does not reinforce the positive association with population growth.

2.C.2 Additional OLS results: time-varying controls

In Section 2.C.1, I use the full city sample from Buringh (2021), for which time-varying controls are not available. This raises concerns about omitted variable bias. To address this, I demonstrate here that including city and country-by-time fixed effects already captures the most important sources of variation, as results remain stable when time-varying controls are added.

TABLE 2.9: OLS Estimates: Effect of Scientific and Literary Academies on City Population.

	Dependent Variable: ln pop in 1500-1900			
	Obs.: **ALL cities** as in Buringh (2021)			
	(1)	(2)	(3)	(4)
ACAD$_{Science}$	0.365**		0.443***	
	(0.178)		(0.319)	
ACAD$_{Literary}$		0.172		0.089
		(0.122)		(0.157)
UNI			0.098*	0.090
			(0.056)	(0.071)
ACAD$_{Science}$**xUNI**			−0.184	
			(0.319)	
ACAD$_{Literary}$**xUNI**				−0.167
				(0.235)
Cities	2096	2096	2096	2096
R^2	0.845	0. 845	0.845	0.845
city FE	YES	YES	YES	YES
countryXtime FE	YES	YES	YES	YES

Note: *p<0.1; **p<0.05; ***p<0.01.
Standard errors clustered at the city level are reported in parentheses.
ACAD$_{\text{Science}}$ indicates the presence of a scientific academy in a given city and time period; **ACAD$_{\text{Literary}}$** indicates the presence of a literary academy.
ACAD$_{\text{Science}}$ × **UNI** and **ACAD$_{\text{Literary}}$** × **UNI** represent interaction terms capturing the joint presence of a university and a scientific or literary academy in the same city and period.
The dependent variable is the natural logarithm of city population, measured at 50-year intervals from 1500 to 1900, based on the dataset in Buringh (2021).
All regressions include city fixed effects and country-by-time fixed effects. Each column represents a separate specification.

This check is possible only for the subset of large cities from Bosker, Buringh, and Van Zanden (2013), for which time-varying characteristics are available until 1800. I first identify relevant variables by examining their temporal variation. Several factors clearly change over time: for example, cities were plundered with differing frequency across years; Bruges and Seville lost direct access to the sea; some cities gained or lost status as a bishopric, archbishopric, or capital. The religious landscape also shifted, as seen in cities like Granada, which ceased to host a madrasa between 1500 and 1600, and in changing exposure to Muslim influence. Measures such as Muslim and Christian urban potential—defined by Bosker, Buringh, and Van Zanden (2013, p.1423) as the distance-weighted population of nearby cities—also evolve over time.

In contrast, many controls remain stable: distances to Rome, Mecca, or Byzantium; soil quality (based on Ramankutty et al. (2002)); elevation and terrain variation; and proximity to rivers or Roman roads. These are effectively accounted for by city fixed effects.

While city FE controls for all time-invariant characteristics, it does not capture time-varying dynamics. However, Table 2.10 shows that adding time-varying controls has little effect: none of the coefficients in Columns 2 and 3 becomes significant or changes sign. Column 1 complements Table 2.8 in the main text and confirms similar trends, supporting the robustness of the baseline specifications.

TABLE 2.10: OLS Estimates for Large Cities (as in Bosker, Buringh, and Van Zanden (2013)).

	ln pop in 1500-1900	ln pop in 1500-1800	
	(1)	(2)	(3)
ACAD	0.248*	0.166	0.173
	(0.133)	(0.129)	(0.136)
UNI	0.033	−0.021	−0.062
	(0.068)	(0.068)	(0.068)
ACADxUNI	−0.128	−0.038	−0.086
	(0.169)	(0.163)	(0.169)
Cities	633	633	633
Adj. R^2	0.823	0.810	0.800
$\Delta time$ controls	NO	YES	NO
$city$ FE	YES	YES	YES
$countryXtime$ FE	YES	YES	YES

Note: *p<0.1; **p<0.05; ***p<0.01.
Standard errors are clustered at the city level and reported in parentheses.
The dependent variable is the natural logarithm of population in large cities over three periods: 1500–1900 (Column 1), and 1500–1800 (Columns 2 and 3).
ACAD indicates the presence of an academy; **UNI** indicates the presence of a university; **ACAD** × **UNI** captures cities hosting both types of institutions in the same period.
City and country-by-time fixed effects are included in all specifications. Column (2) additionally controls for time-varying characteristics, including: direct sea access, presence of a bishop or archbishop, capital city status, presence of a madrasa, number of times plundered, Muslim or Christian urban potential, and whether the city was predominantly Muslim.

2.C.3 Dynamic TWFE results

This section presents the main results from the dynamic TWFE estimations, addressing two key identifying assumptions: parallel trends and no anticipation effects. Section 2.7.2 further considers the SUTVA assumption and investigates possible inbound spatial spillovers (Berkes & Nencka, 2021; Butts, 2021).

Figure 2.14 displays the main event study, analyzing the effects of academy creation between 1500 and 1900 in 50-year intervals. The assumptions of no pre-trends and no anticipation are met.[54] Results show a significant positive effect emerging 100 years after the academy's founding, persisting throughout the last century of the period studied. This implies that, in the long run, cities with academies experience faster population growth than comparable cities without them.

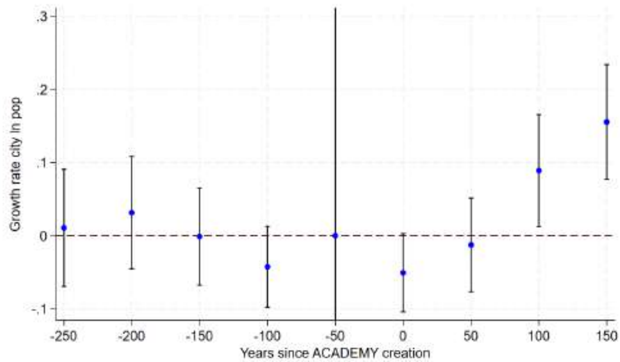


FIGURE 2.14: Event Study: Academy Creation.

This figure shows the estimated effects of academy creation on city population from 1500 to 1900, based on an event-study specification.
Note: The control group is cities that never established an academy. The dependent variable is the logarithmic city population growth rate. The sample includes 2,023 city-level clusters. The within R^2 is 0.315.

To examine the interaction between universities and academies, I restrict the sample to cities that hosted a university at any point. Since universities were typically established before academies, this subset offers a more comparable group of cities and allows for a cleaner investigation of the interaction term.[55] However, this reduced sample has lower statistical power. Figure 2.15 shows that while parallel trends hold, no significant effect follows academy creation in university cities.

Next, I explore heterogeneity by field of study, longevity, and size of the academies.

[54]Testing for pre-trends offers only a partial check of the parallel trends assumption.

[55]Only in Nancy (France) was the academy created clearly before the university. In two other cases—Halle (Germany) and Pau (France)—the academy slightly precedes the university but within the same 50-year period.

I first distinguish between scientific and literary academies based on member composition. An academy is classified as scientific if over 50% of its members focus on science, applied science, or medicine (43 academies); literary academies include those where over 50% of members engage in literature, history, theology, law, or social sciences (38 academies).

Figure 2.16, Panel (a), shows a strong and lasting positive effect from scientific academies, with population growth increasing by 23% (p = 0.000) 150 years after establishment. In contrast, Panel (b) shows no significant effect from literary academies. A mild pre-trend is present but in the opposite direction of the post-treatment effect, suggesting any bias would attenuate the observed results.

I also analyze large academies (more than 30 members; 62 cases) and long-lasting academies (active over 30 years; 68 cases). Both groups show similar patterns: a significant positive effect appears only after 150 years.[56]

These findings emphasize the importance of scholarly focus: it is not size or longevity that drives outcomes, but the scientific orientation of the academy. Institutions centered on experimental and applied inquiry—rather than literary or humanistic pursuits—have the strongest and most persistent impact on long-run urban growth.

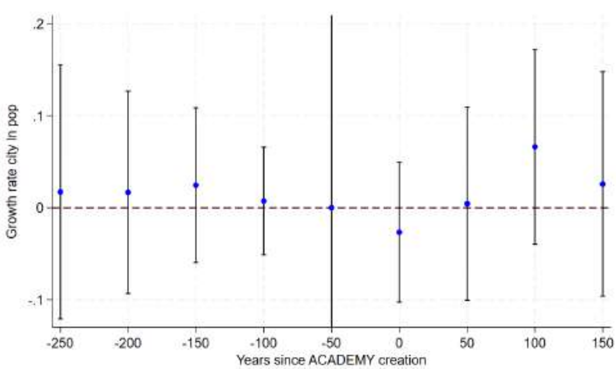
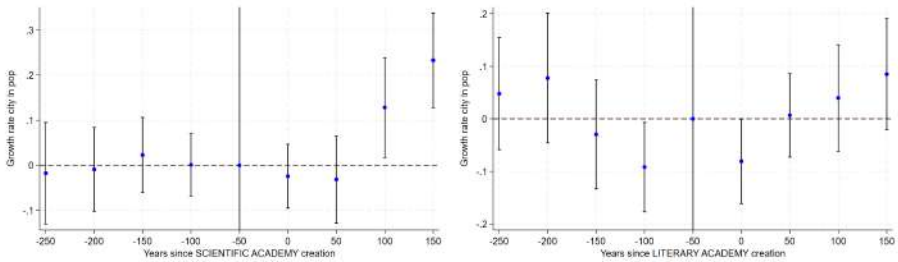


FIGURE 2.15: Event Study: Academy Creation in University cities.

This figure shows the estimated effects of academy creation on city population from 1500 to 1900 in cities that hosted a university at least once, based on an event-study specification.
Note: The control group is university cities that never established an academy. The dependent variable is the logarithmic city population growth rate. The sample includes 149 city-level clusters. The within R^2 is 0.528.

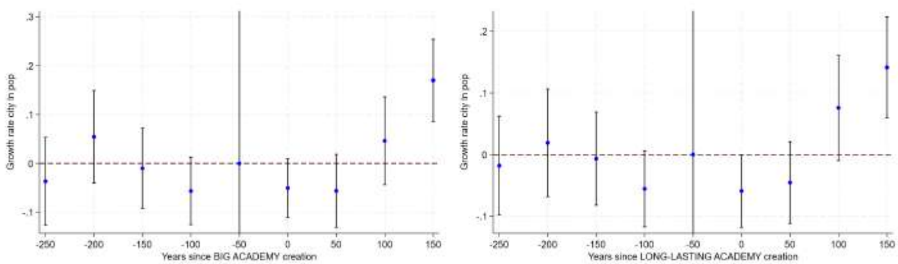
[56]See Figures 2.17a and 2.17b.



(A) (B)

FIGURE 2.16: Event Study: Academy Creation by Field.

This figure shows the estimated effects of academy creation on city population from 1500 to 1900, based on an event-study specification. Panel (a) refers to scientific academies; panel (b) to literary academies.
Note: The control group is never-treated cities. The dependent variable is the logarithmic city population growth rate. The sample includes 2,023 city-level clusters. The within R^2 is 0.315 for both panels.



(A) (B)

FIGURE 2.17: Event Study: Academy Creation by Size and Duration.

This figure shows the estimated effects of academy creation on city population from 1500 to 1900, based on an event-study specification. Panel (a) refers to large academies; panel (b) to long-lasting academies.
Note: The control group is never-treated cities. The dependent variable is the logarithmic city population growth rate. The sample includes 2,023 city-level clusters. The within R^2 is 0.315 for both panels.

2.C.4 Static TWFE

In this section, I estimate a traditional static Two-Way Fixed Effects (TWFE) specification as follows:

$$lnPOP_{ct} = \beta_0 + \beta_1 EVENT_c \text{x} Post_{ct} + \mu_c + \psi_s \lambda_t + \epsilon_{ct} \tag{2.2}$$

Here, $lnPOP_{ct}$ denotes the logarithm of the population of city c at time t, which is the outcome variable. $Post_{ct}$ is an indicator equal to 1 in all periods following the creation of an academy in city c, while $EVENT_c$ identifies whether the event occurred in the city. The interaction term $EVENT_c\text{x}Post_{ct}$ captures the treatment effect. The coefficient of interest, β_1, measures the average change in population size after the establishment of an academy, relative to cities without academies.

The model includes city fixed effects μ_c and country-by-time fixed effects $\psi_s \lambda_t$ to control for unobserved heterogeneity across cities and over time.

While the static TWFE model provides a straightforward interpretation, it tends to deliver conservative estimates, as it downweights long-term effects in favor of short-term ones. Moreover, the assumption of homogeneous treatment effects across time and space is unlikely to hold in this setting, limiting the credibility of causal identification through this approach.

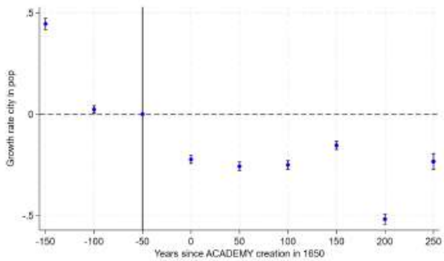
I report these static estimates for completeness. However, the main analysis relies on more robust dynamic TWFE models and advanced difference-in-differences estimators, which are discussed in the main text.

TABLE 2.11: Static Two-Way Fixed Effects: Academy Creation.

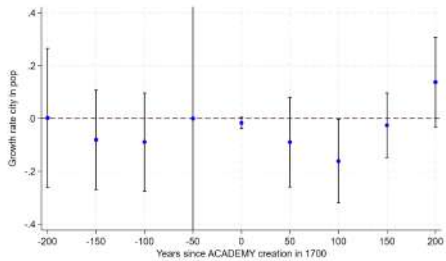
	ln pop in 1500-1900		
	(1)	(2)	(3)
ACAD x Post	0.255***		
	(0.072)		
ACAD$_{science}$ **x Post**		0.277***	
		(0.096)	
ACAD$_{literary}$ **x Post**			0.217**
			(0.106)
Constant	2.969***	2.9731***	2.975***
	(0.013)	(0.013)	(0.013)
Obs.	18207	18207	18207
within R^2	0.725	0.724	0.724
$city$ FE	YES	YES	YES
$countryXtime$ FE	YES	YES	YES

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors are clustered at the city level in parenthesis. The dependent variable is the logarithm of city population. All regressions include city and country-by-time fixed effects. This table reports estimates from static two-way fixed effects regressions, examining the association between academy creation and city population from 1500 to 1900. Column (1) shows the effect of any academy; column (2) restricts to scientific academies; column (3) to literary ones.

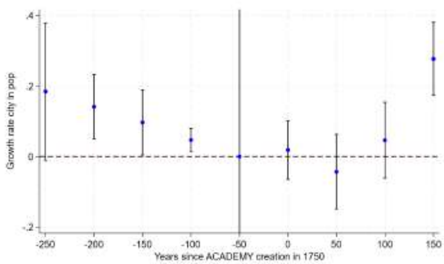
2.C.5 2x2 Event Studies



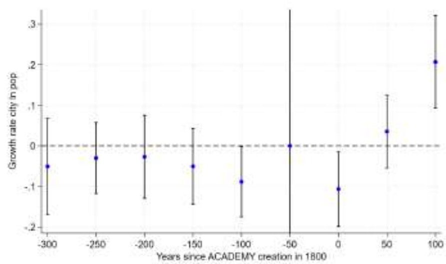
(A)



(B)



(C)



(D)

FIGURE 2.18: Event Study: Academy Creation by Period.

This figure shows the estimated effects of academy creation on city population using a 2x2 event-study design for four selected periods: (a) 1650 — only the Investiganti Academy in Naples was founded; (b) 1700 — 14 academies were established; (c) 1750 — 26 academies were created; (d) 1800 — 40 academies were opened.

Note: The control group is never-treated and cities treated in a different period of time.The dependent variable is the logarithmic city population growth rate. The sample includes 2,023 city-level clusters. The within R^2 is (a) 0.314, (b) 0.314, (c) 0.315, (d) 0.316.

2.D Alternative DID estimators

2.D.1 $CSDID$ results

As in Sun and Abraham (2021), Callaway and Sant'Anna (2021) focuses on estimating average treatment effects by cohort—groups of cities g that experience the creation of an academy for the first time at time t. The $CSDID$ estimator developed by Callaway and Sant'Anna (2021) is methodologically close to the IW estimator from Sun and Abraham (2021), which I use for the main results in Section 2.5.1.

Compared to Sun and Abraham (2021), Callaway and Sant'Anna (2021) offers greater flexibility, allowing for various aggregation methods. However, these features are less relevant in my setting. Instead, I rely on the event-study structure from Sun and Abraham (2021) to ensure balanced timing across cohorts. Still, due to the methodological similarities, I report results using the $CSDID$ estimator as a robustness check.

Figure 2.19 presents the dynamic treatment effects from creating an academy, estimated using a balanced panel with five pre-treatment and three post-treatment periods. The results mirror those in the main analysis: effects become statistically significant at the 90% level in later periods, consistent with the long-run impact detected by Sun and Abraham (2021). The pre-treatment coefficient is small and not statistically significant (-0.7%, p-value: 0.69), providing reassurance against differential pre-trends. However, the average treatment effect on the treated (ATT) is only marginally insignificant (4.5%, p-value: 0.11), suggesting that the overall effects may be modest when using this estimator.

An additional feature of $CSDID$ is that it allows for interpretation of effects by calendar time. For example, in Naples—the only city to receive an academy in 1650—the estimated population growth rate is 22% lower (p-value: 0.000) between 1650 and 1700 compared to a counterfactual without an academy. A significant positive effect (21%, p-value: 0.000) is only observed after 1900.

Figure 2.20 shows the dynamic effects of academy creation in cities that hosted a university at least once. As in Figure 2.4, there are no significant pre-trends (0.35%, p-value: 0.91) or post-treatment effects (-2.4%, p-value: 0.51), confirming the absence of interaction effects on population growth.

Figure 2.21 displays the dynamics by type of academy. Scientific academies are associated with a delayed but positive effect on city population growth: after 100 years, growth is 12.5% higher on average (p-value: 0.059), with no pre-trends (-0.2%, p-value: 0.93). The ATT is slightly insignificant (6.6%, p-value: 0.12), but the final calendar year (1900) shows a significant 29% increase (p-value: 0.00).

Conversely, literary academies (Figure 2.21b) are associated with short-term slow down in growth: an 8% decrease immediately after the creation of the

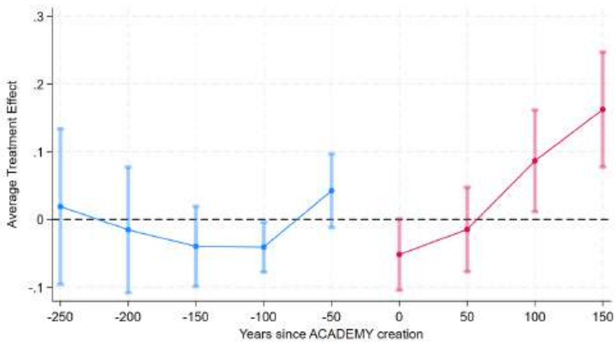


FIGURE 2.19: Academy event using Callaway and Sant'Anna (2021).

This figure shows the dynamic average effects of creating an academy on city population growth between 1500 and 1900, estimated using the method of Callaway and Sant'Anna (2021).
Note: The control group is cities that never established an academy. The dependent variable is the logarithmic city population growth rate.

academy. There are no significant pre-trends (−1.2%, p-value: 0.60) which is reassuring also no joint post-treatment effect. In addition, three post-treatment calendar periods show statistically significant negative effects, only after 1900 does the effect turn positive (17%, p-value: 0.005).

Finally, Figures 2.21c and 2.21d examine long-lasting academies (over 30 years) and large academies (over 30 members). Their dynamic patterns are similar to the overall case. While no significant pre-trends are detected, there is also no statistically significant average post-treatment effect.

2.D.2 DID_l results

The estimator proposed by De Chaisemartin and d'Haultfoeuille (2024), denoted as DID_l, measures the effect of creating an academy or university exactly l periods ago for the first time. It compares cities undergoing the initial creation of a higher education institution at time t with those that have not yet received such a treatment. This method captures the delayed effects of institutional creation and is particularly suited for assessing long-difference dynamics. A key innovation of DID_l is its ability to handle continuous treatments.[57]

[57] Earlier drafts of this project relied on the estimator from De Chaisemartin and d'Haultfoeuille (2022), which yielded nearly identical results. I also compared findings with De Chaisemartin and d'Haultfoeuille (2020), which estimates the instantaneous treatment effect by focusing solely on cities whose treatment status changes in the current period. That approach excludes cities that

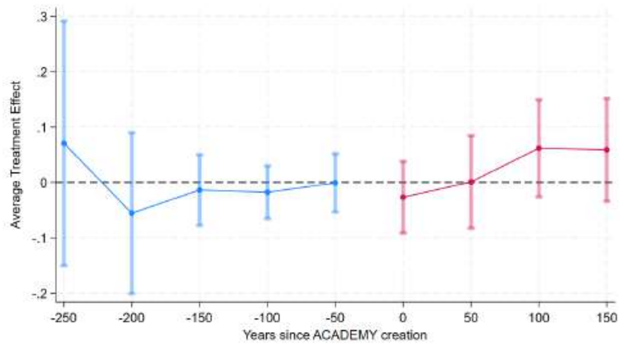


FIGURE 2.20: Academy event in university cities using Callaway and Sant'Anna (2021).

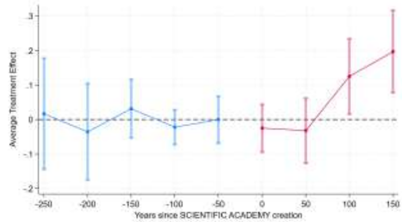
This figure shows the dynamic average effects of creating an academy between 1500 and 1900 in cities that hosted a university at least once, estimated using Callaway and Sant'Anna (2021). *Note:* The control group is university cities that never established an academy. The dependent variable is the logarithmic city population growth rate.

In what follows, I present results from the DID_l estimator. Since it does not allow for more leads than lags, I include a balanced specification with three leads and three lags. This setup provides a more nuanced perspective on the timing and persistence of treatment effects.

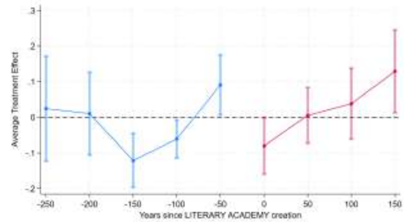
Figure 2.22a reports the estimated impact of academy creation between 1500 and 1900. The dynamic closely mirrors the baseline: a significant positive effect emerges roughly 100 years after the academy's establishment. The same pattern holds for scientific academies, which increase city population growth by 12.5% after a century, as shown in Figure 2.23. In contrast, literary academies show no statistically significant effect on population growth.

Finally, Figure 2.22b illustrates the dynamics for cities that had at least one university. These results reinforce previous findings—there is no significant interaction effect between universities and academies on city growth.

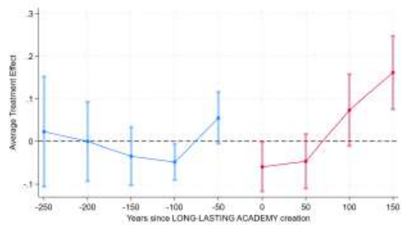
have always or never been treated since 1500. While results were again broadly consistent, placebo estimates varied slightly due to differences in treatment status definitions.



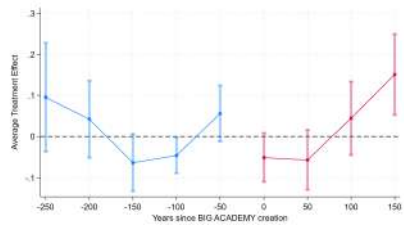
(A)



(B)



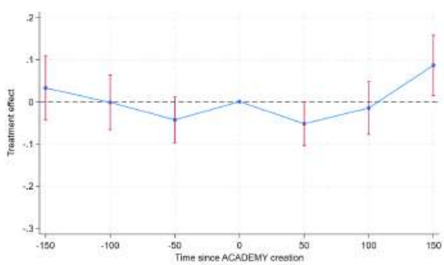
(C)



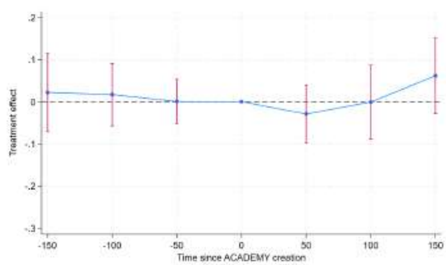
(D)

FIGURE 2.21: Academy event by field, size, and length using Callaway and Sant'Anna (2021).

This figure shows the dynamic average effects of creating (a) a scientific academy, (b) a literary academy, (c) a long-lasting academy (active more than 30 years), and (d) a big academy (with more than 30 members) between 1500 and 1900, estimated using Callaway and Sant'Anna (2021).
Note: The control group is never-treated cities. The dependent variable is the logarithmic city population growth rate.



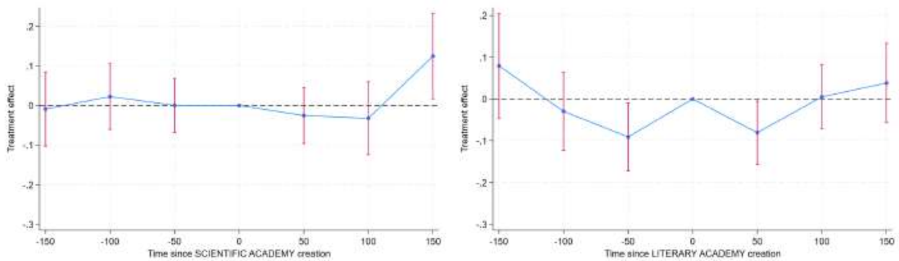
(A) Academy event.



(B) Academy event in university cities.

FIGURE 2.22: Academy event using DID_l (De Chaisemartin & d'Haultfoeuille, 2024)

This figure shows the effect of creating an academy estimated using DID_l (De Chaisemartin & d'Haultfoeuille, 2024) between 1500 and 1900.
Note: The control group is (a) cities that never established an academy (b) university cities that never established an academy. The dependent variable is the logarithmic city population growth rate.



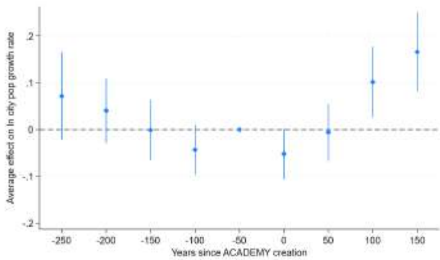
(A) (B)

FIGURE 2.23: Academy event by field, size, and length using De Chaisemartin and d'Haultfoeuille (2024).

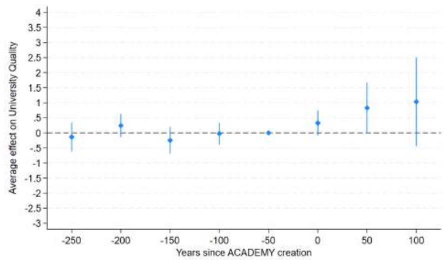
This figure shows the effect estimated using DID_l (De Chaisemartin & d'Haultfoeuille, 2024) of creating (a) a scientific academy and (b) a literary academy between 1500 and 1900.
Note: The control group is never-treated cities. The dependent variable is the logarithmic city population growth rate.

2.E Robustness checks

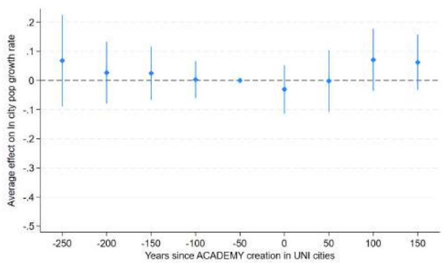
2.E.1 Sensitivity analyses: leave-one-out



(A)



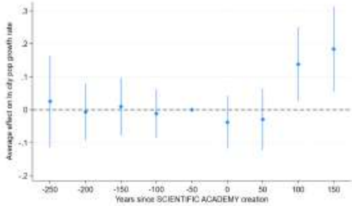
(B)



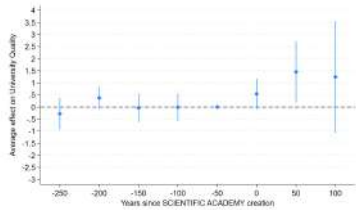
(C)

FIGURE 2.24: Academy events leaving London out.

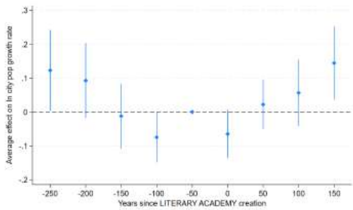
This figure shows the effect of creating (a) an academy on logarithmic city population growth rate, (b) an academy on university quality, and (c) an academy in cities that hosted a university at least once on logarithmic city population growth rate; estimated using Sun and Abraham (2021). *Note:* The control group is cities that never established an academy. London is excluded from the sample.



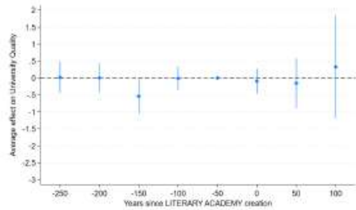
(A)



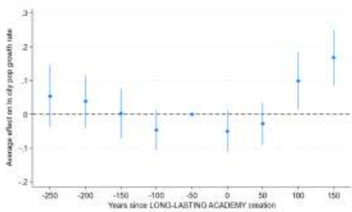
(B)



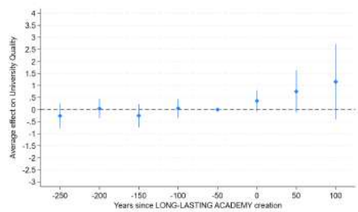
(C)



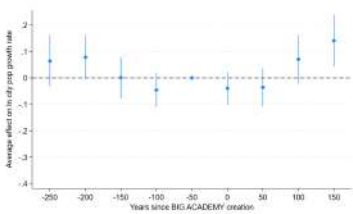
(D)



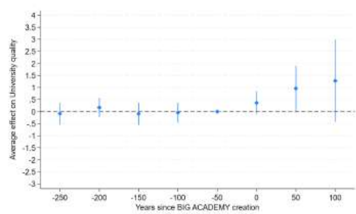
(E)



(F)



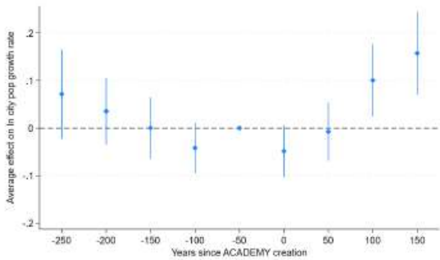
(G)



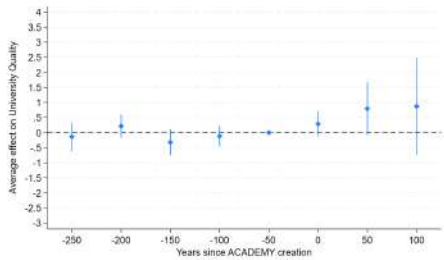
(H)

FIGURE 2.25: Academy events by field, size, and length leaving London out.

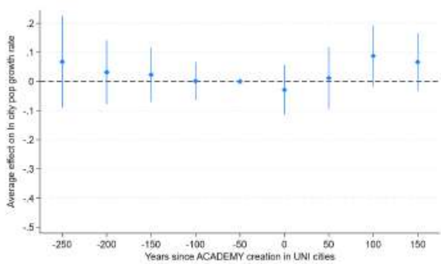
Effect of creating (a - b) a scientific academy, (c - d) a literary academy, (e - f) a long-lasting academy, and (g - h) a big academy; estimated using Sun and Abraham (2021).
Note: The control group is never-treated cities. The dependent variable is logarithmic city population growth rate in the left column, and university quality in the right column. London is excluded from the sample.



(A)



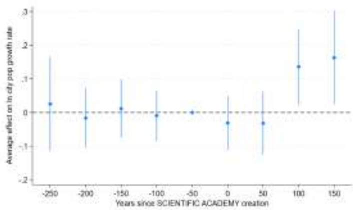
(B)



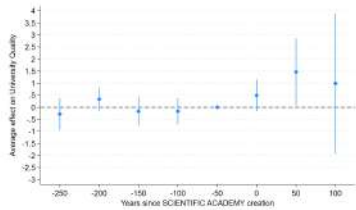
(C)

FIGURE 2.26: Academy events leaving Paris out.

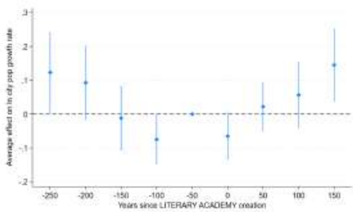
This figure shows the effect of creating (a) an academy on logarithmic city population growth rate, (b) an academy on university quality, and (c) an academy in cities that hosted a university at least once on logarithmic city population growth rate; estimated using Sun and Abraham (2021). *Note:* The control group is cities that never established an academy. Paris is excluded from the sample.



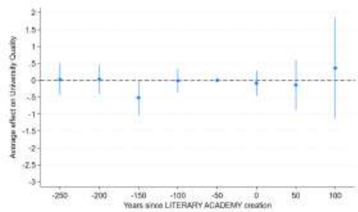
(A)



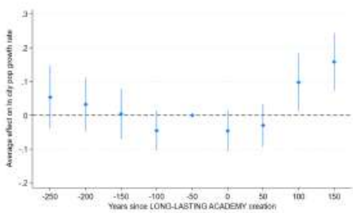
(B)



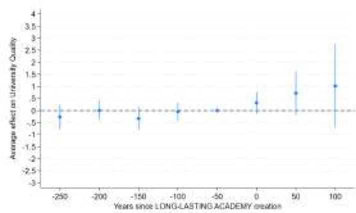
(C)



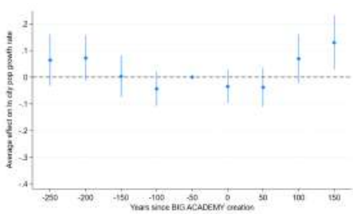
(D)



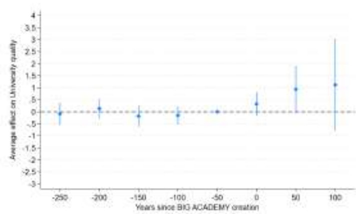
(E)



(F)



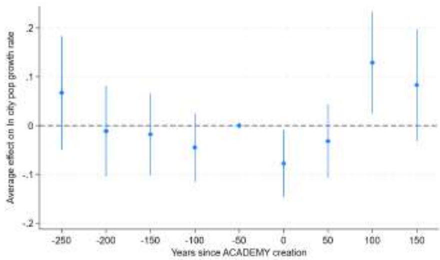
(G)



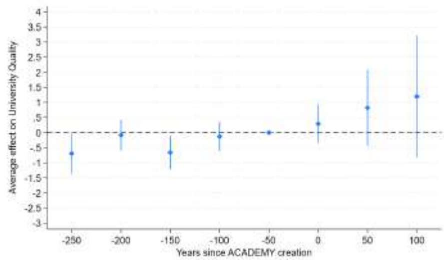
(H)

FIGURE 2.27: Academy events by field, size, length leaving Paris out.

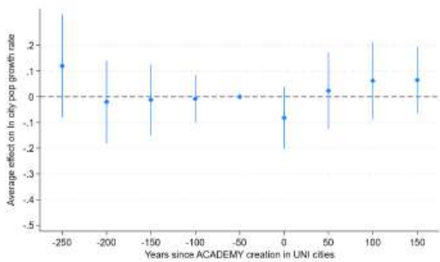
Effect of creating (a - b) a scientific academy, (c - d) a literary academy, (e - f) a long-lasting academy, and (g - h) a big academy; estimated using Sun and Abraham (2021).
Note: The control group is never-treated cities. The dependent variable is logarithmic city population growth rate in the left column, and university quality in the right column. Paris is excluded from the sample.



(A)



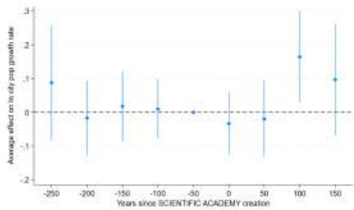
(B)



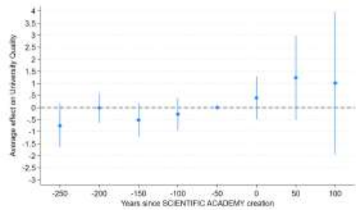
(C)

FIGURE 2.28: Academy events leaving France out.

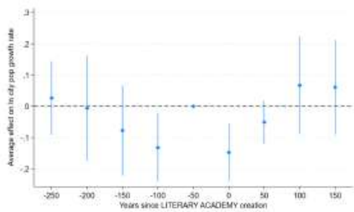
This figure shows the effect of creating (a) an academy on logarithmic city population growth rate, (b) an academy on university quality, and (c) an academy in cities that hosted a university at least once on logarithmic city population growth rate; estimated using Sun and Abraham (2021). *Note:* The control group is cities that never established an academy. France is excluded from the sample.



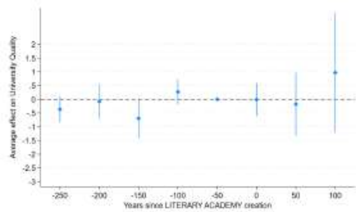
(A)



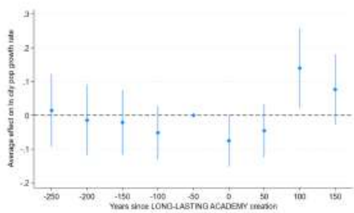
(B)



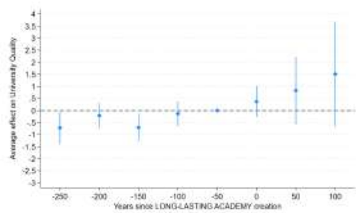
(C)



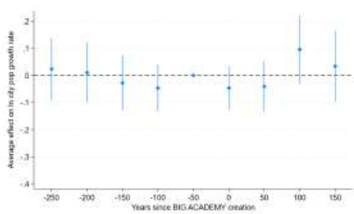
(D)



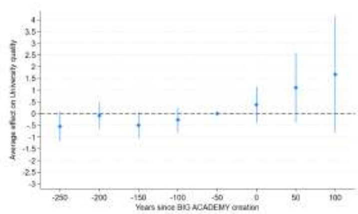
(E)



(F)



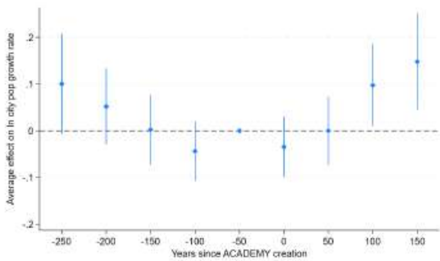
(G)



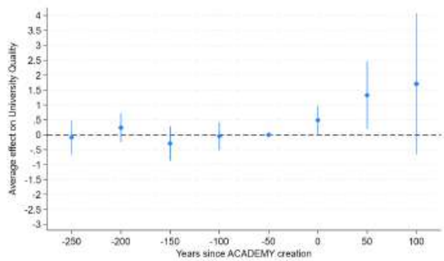
(H)

FIGURE 2.29: Academy events by field, size, length leaving France out.

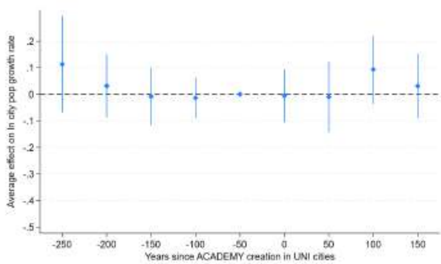
Effect of creating (a - b) a scientific academy, (c - d) a literary academy, (e - f) a long-lasting academy, and (g - h) a big academy; estimated using Sun and Abraham (2021).
Note: The control group is never-treated cities. The dependent variable is logarithmic city population growth rate in the left column, and university quality in the right column. France is excluded from the sample.



(A)



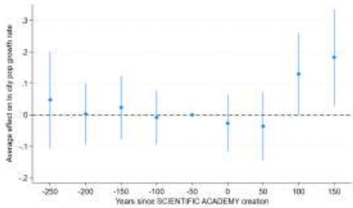
(B)



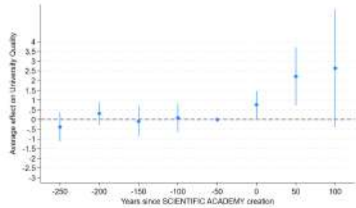
(C)

FIGURE 2.30: Academy events leaving Italy out.

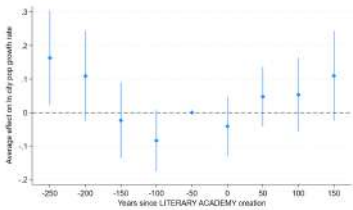
This figure shows the effect of creating (a) an academy on logarithmic city population growth rate, (b) an academy on university quality, and (c) an academy in cities that hosted a university at least once on logarithmic city population growth rate; estimated using Sun and Abraham (2021). *Note:* The control group is cities that never established an academy. Italy is excluded from the sample.



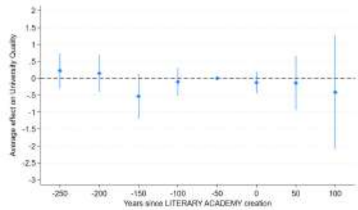
(A)



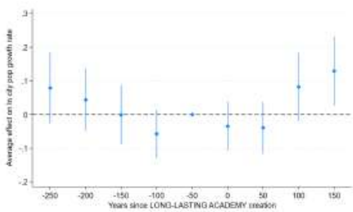
(B)



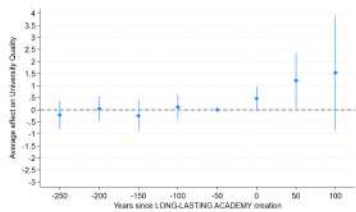
(C)



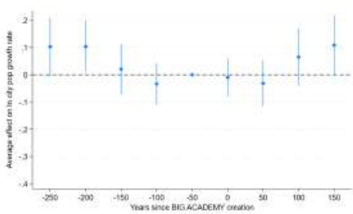
(D)



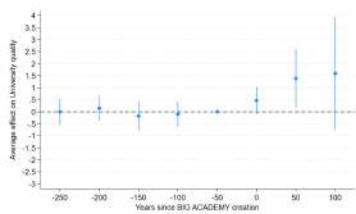
(E)



(F)



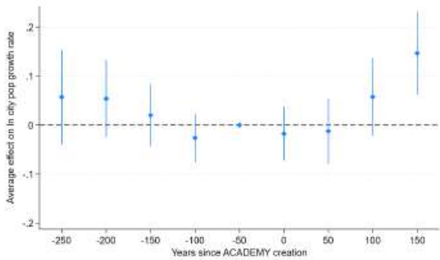
(G)



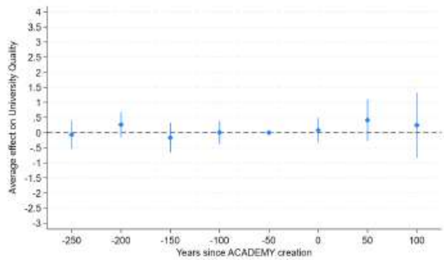
(H)

FIGURE 2.31: Academy events by field, size, length leaving Italy out.

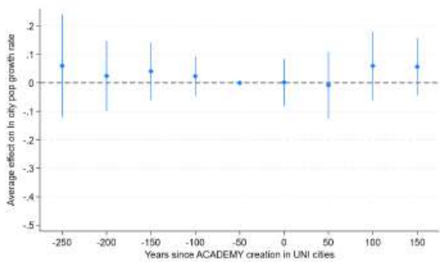
Effect of creating (a - b) a scientific academy, (c - d) a literary academy, (e - f) a long-lasting academy, and (g - h) a big academy; estimated using Sun and Abraham (2021).
Note: The control group is never-treated cities. The dependent variable is logarithmic city population growth rate in the left column, and university quality in the right column. Italy is excluded from the sample.



(A)



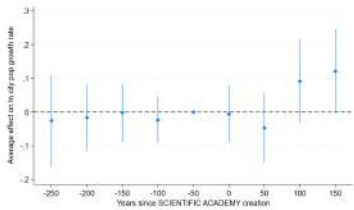
(B)



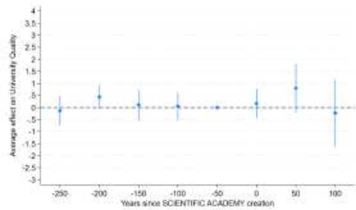
(C)

FIGURE 2.32: Academy events leaving Germany out.

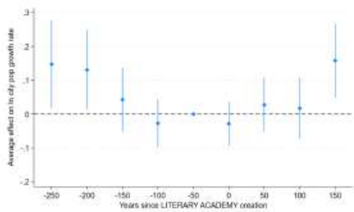
This figure shows the effect of creating (a) an academy on logarithmic city population growth rate, (b) an academy on university quality, and (c) an academy in cities that hosted a university at least once on logarithmic city population growth rate; estimated using Sun and Abraham (2021). *Note:* The control group is cities that never established an academy. Germany is excluded from the sample.



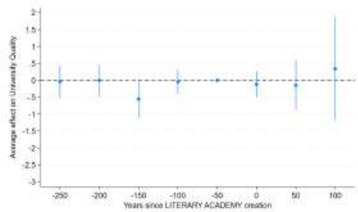
(A)



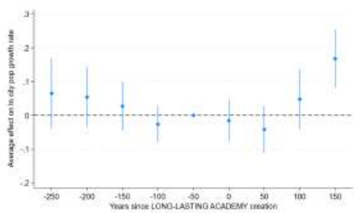
(B)



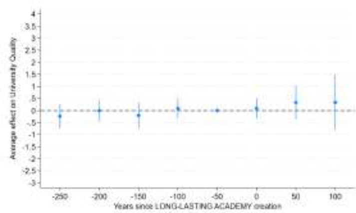
(C)



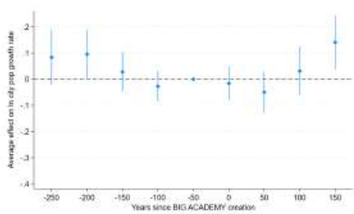
(D)



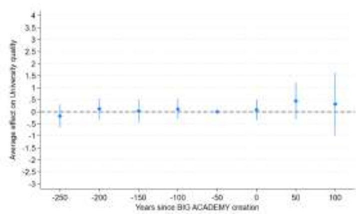
(E)



(F)



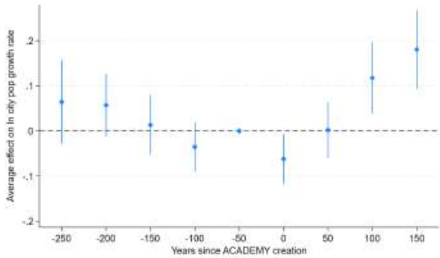
(G)



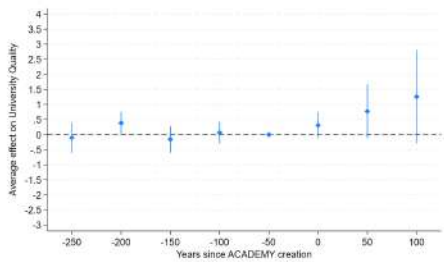
(H)

FIGURE 2.33: Academy events by field, size, length leaving Germany out.

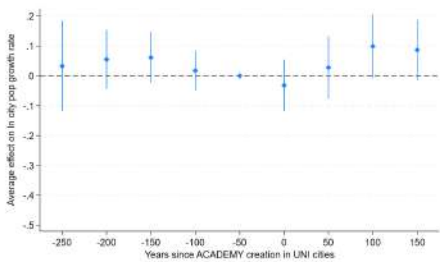
Effect of creating (a - b) a scientific academy, (c - d) a literary academy, (e - f) a long-lasting academy, and (g - h) a big academy; estimated using Sun and Abraham (2021).
Note: The control group is never-treated cities. The dependent variable is logarithmic city population growth rate in the left column, and university quality in the right column. Germany is excluded from the sample.



(A)



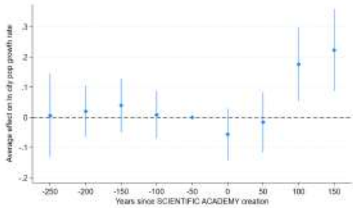
(B)



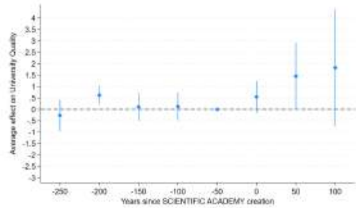
(C)

FIGURE 2.34: Academy events leaving Great Britain out.

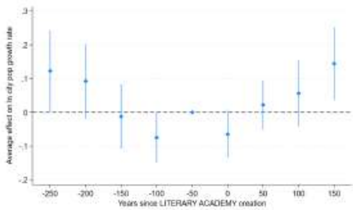
This figure shows the effect of creating (a) an academy on logarithmic city population growth rate, (b) an academy on university quality, and (c) an academy in cities that hosted a university at least once on logarithmic city population growth rate; estimated using Sun and Abraham (2021). *Note:* The control group is cities that never established an academy. Great Britain is excluded from the sample.



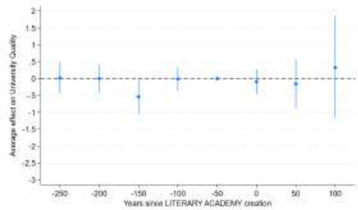
(A)



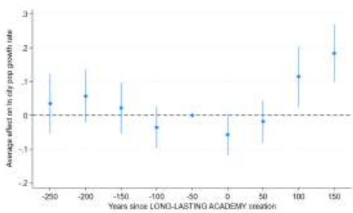
(B)



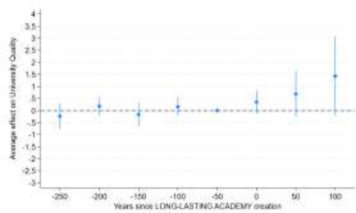
(C)



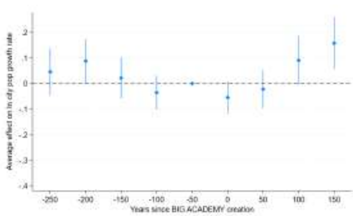
(D)



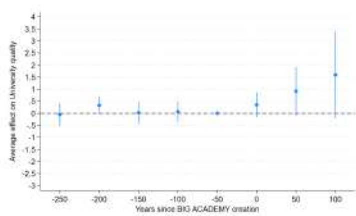
(E)



(F)



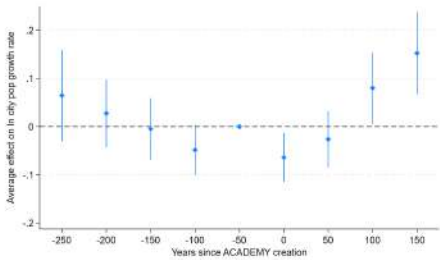
(G)



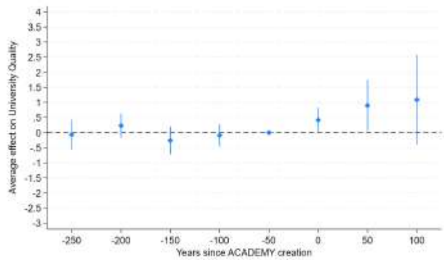
(H)

FIGURE 2.35: Academy events by filed, size, length leaving Great Britain out.

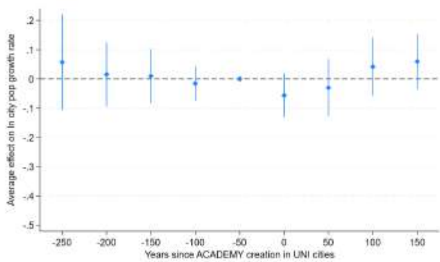
Effect of creating (a - b) a scientific academy, (c - d) a literary academy, (e - f) a long-lasting academy, and (g - h) a big academy; estimated using Sun and Abraham (2021).
Note: The control group is never-treated cities. The dependent variable is logarithmic city population growth rate in the left column, and university quality in the right column. Great Britain is excluded from the sample.



(A)



(B)



(C)

FIGURE 2.36: Academy events leaving Spain out.

This figure shows the effect of creating (a) an academy on logarithmic city population growth rate, (b) an academy on university quality, and (c) an academy in cities that hosted a university at least once on logarithmic city population growth rate; estimated using Sun and Abraham (2021). *Note:* The control group is cities that never established an academy. Spain is excluded from the sample.

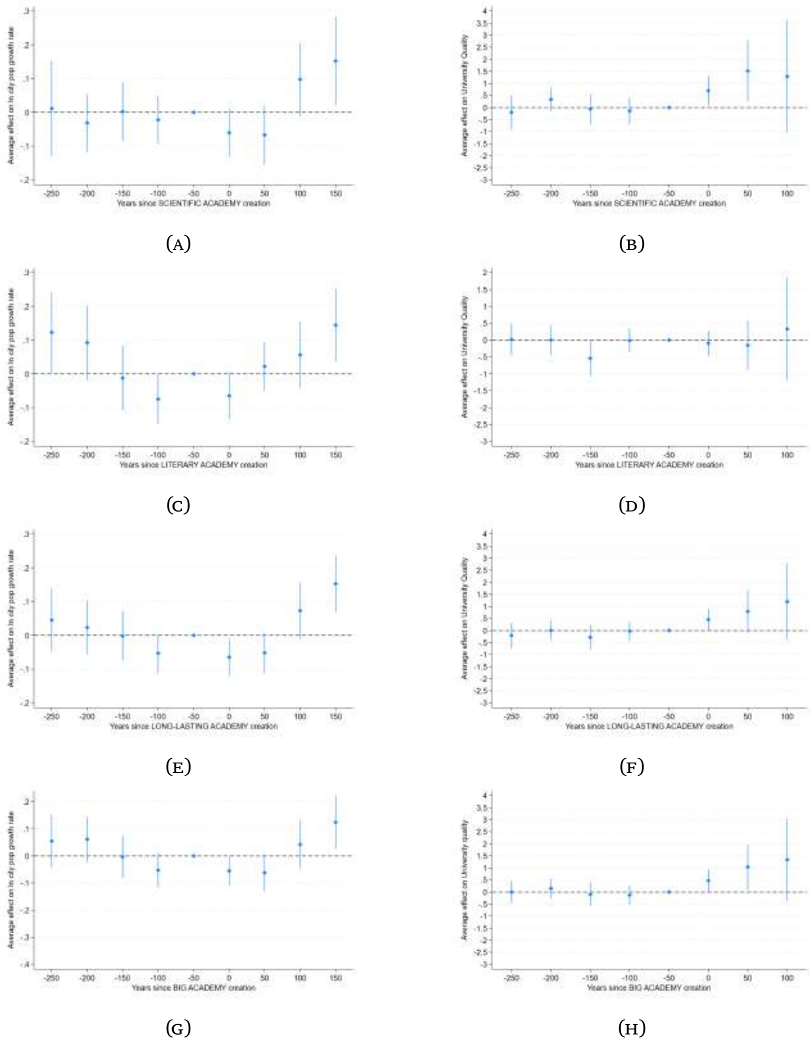
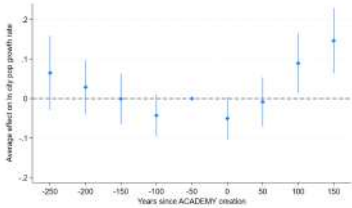


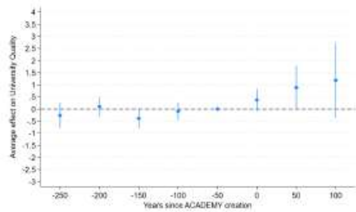
FIGURE 2.37: Academy events by field, size, length leaving Spain out.

Effect of creating (a - b) a scientific academy, (c - d) a literary academy, (e - f) a long-lasting academy, and (g - h) a big academy; estimated using Sun and Abraham (2021).
Note: The control group is never-treated cities. The dependent variable is logarithmic city population growth rate in the left column, and university quality in the right column. Spain is excluded from the sample.

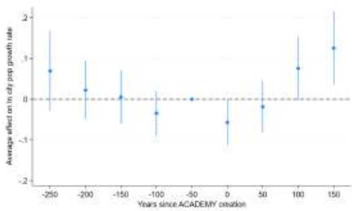
2.E.2 Local effects



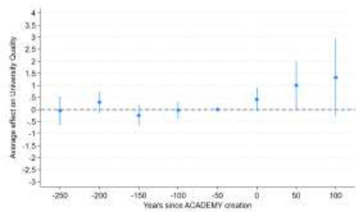
(A)



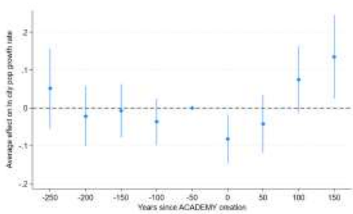
(B)



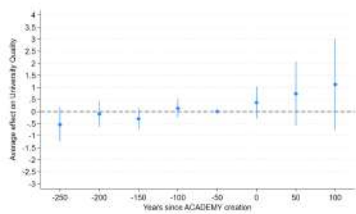
(C)



(D)



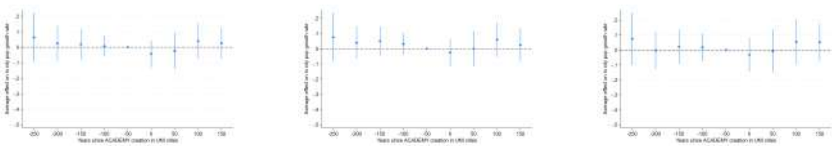
(E)



(F)

FIGURE 2.38: Test for local effects of academy events.

Effect of creating an academy excluding cities (a - b) within 50 km, (c - d) within 100 km, and (e - f) within 150 km from the seat of the academy; estimated with Sun and Abraham (2021).
Note: The control group is cities that never established an academy. The dependent variable is logarithmic city population growth rate in the left column, and university quality in the right column.

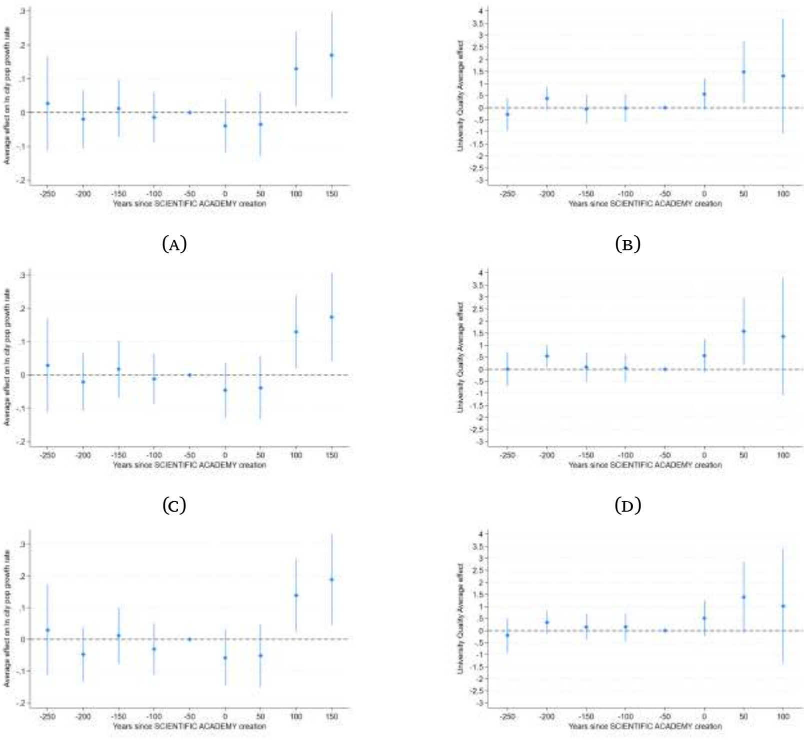


(A) (B) (C)

FIGURE 2.39: Test for local effects of academy events in university cities.

Effect of creating an academy in cities that ever had a university, excluding cities (a) within 50 km, (b) within 100 km, and (c) within 150 km from the seat of the academy; estimated with Sun and Abraham (2021).

Note: The control group is university cities that never established an academy. The dependent variable is the logarithmic city population growth rate.



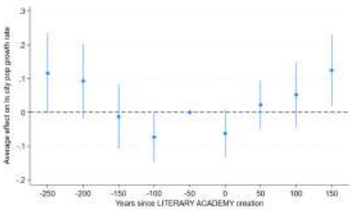
(A) (B)

(C) (D)

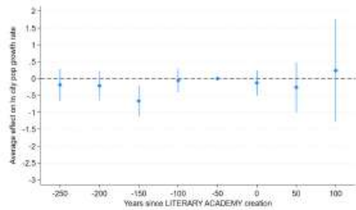
(E) (F)

FIGURE 2.40: Test for local effects of scientific academy events.

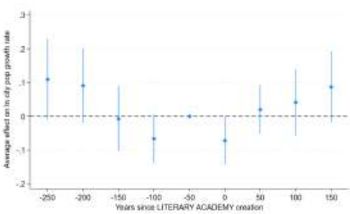
Effect of creating a scientific academy excluding cities (a - b) within 50 km, (c - d) within 100, (e - f) within 150km from the seat of the academy; estimated with Sun and Abraham (2021).
Note: The control group is cities that never established a scientific academy. The dependent variable is the logarithmic city population growth rate on the left column, and quality of universities on the right column.



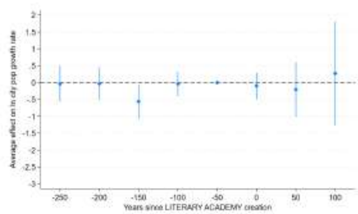
(A)



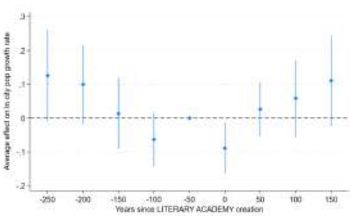
(B)



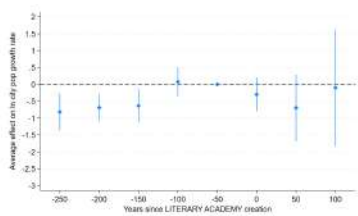
(C)



(D)



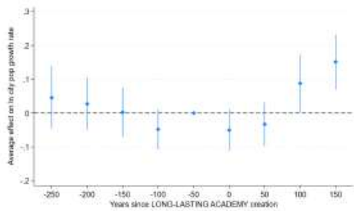
(E)



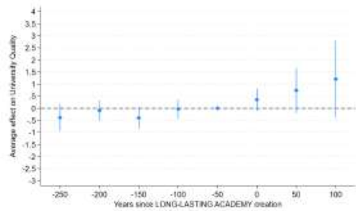
(F)

FIGURE 2.41: Test for local effects of literary academy events.

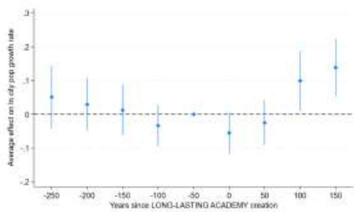
Effect of creating a literary academy excluding cities (a - b) within 50 km, (c - d) within 100, (e - f) within 150km from the seat of the academy; estimated with Sun and Abraham (2021).
Note: The control group is cities that never established a literary academy. The dependent variable is the logarithmic city population growth rate on the left column, and quality of universities on the right column.



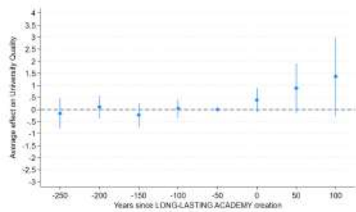
(A)



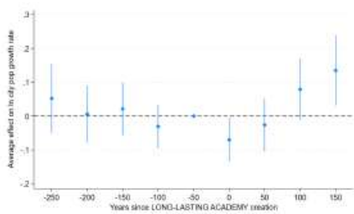
(B)



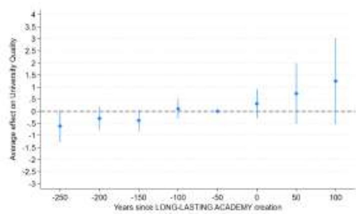
(C)



(D)



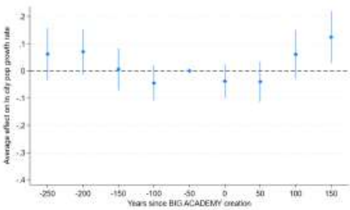
(E)



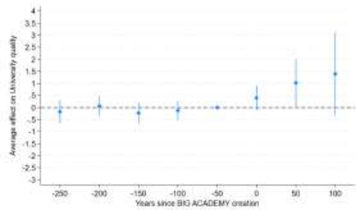
(F)

FIGURE 2.42: Test for local effect of long-lasting academy events.

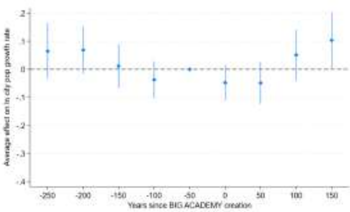
Effect of creating a long-lasting academy excluding cities (a - b) within 50 km, (c - d) within 100, (e - f) within 150km from the seat of the academy; estimated with Sun and Abraham (2021).
Note: The control group is cities that never established a long-lasting academy. The dependent variable is the logarithmic city population growth rate on the left column, and quality of universities on the right column.



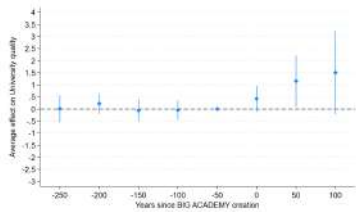
(A)



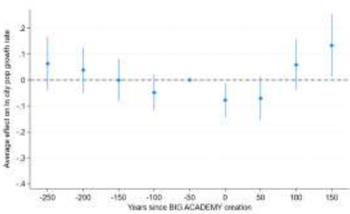
(B)



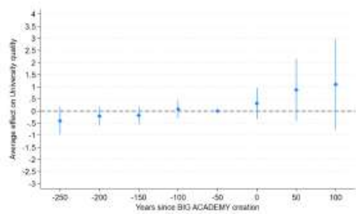
(C)



(D)



(E)

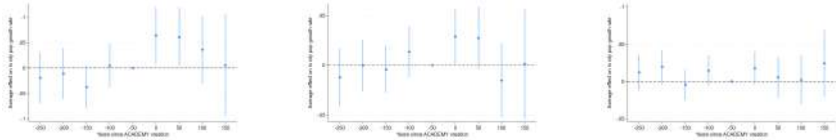


(F)

FIGURE 2.43: Test for local effects of big academy events.

Effect of creating a big academy excluding cities (a - b) within 50 km, (c - d) within 100, (e - f) within 150km from the seat of the academy; estimated with Sun and Abraham (2021).
Note: The control group is cities that never established a big academy. The dependent variable is the logarithmic city population growth rate on the left column, and quality of universities on the right column.

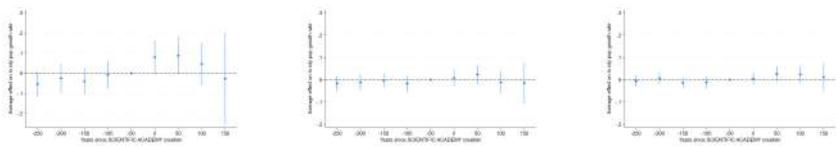
2.E.3 Spillover effects



(A) (B) (C)

FIGURE 2.44: Test for spillover effects of academy events.

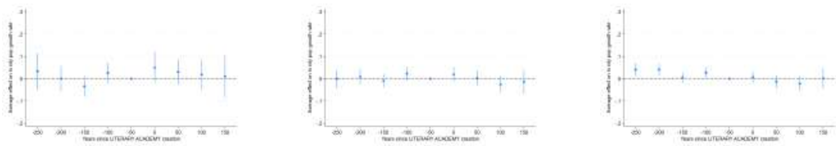
Effect of creating an academy in cities (a) within 25 km from the hosting cities, excluding the hosting city, (b) between 25 km and 50 km from the hosting cities, excluding hosting cities and within the 0–25 km "donut," (c) between 50 km and 75 km from the hosting cities, excluding hosting cities and within the 25–50 km "donut"; estimated with Sun and Abraham (2021).
Note: Control group includes cities that never established an academy and further away than 25 km, 50 km, and 75 km, respectively. Dependent variable is the logarithmic city population growth rate.



(A) (B) (C)

FIGURE 2.45: Test for spillover effects of scientific academy events.

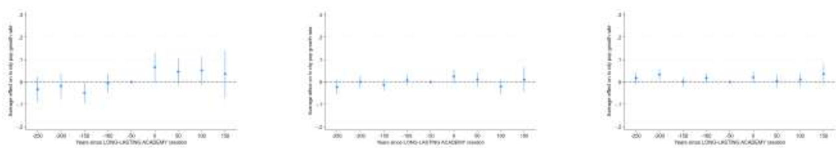
Effect of creating a scientific academy in cities (a) within 25 km from the hosting cities, excluding the hosting city, (b) between 25 km and 50 km from the hosting cities, excluding hosting cities and within the 0–25 km "donut," (c) between 50 km and 75 km from the hosting cities, excluding hosting cities and within the 25–50 km "donut"; estimated with Sun and Abraham (2021).
Note: Control group includes cities that never established a scientific academy and further away than 25 km, 50 km, and 75 km, respectively. Dependent variable is the logarithmic city population growth rate.



(A) (B) (C)

FIGURE 2.46: Test for spillover effects of literary academy events.

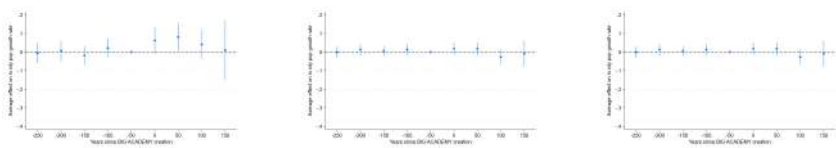
Effect of creating a literary academy in cities (a) within 25 km from the hosting cities, excluding the hosting city, (b) between 25 km and 50 km from the hosting cities, excluding hosting cities and within the 0–25 km "donut," (c) between 50 km and 75 km from the hosting cities, excluding hosting cities and within the 25–50 km "donut"; estimated with Sun and Abraham (2021).
Note: Control group includes cities that never established a literary academy and further away than 25 km, 50 km, and 75 km, respectively. Dependent variable is the logarithmic city population growth rate.



(A) (B) (C)

FIGURE 2.47: Test for spillover effects of long-lasting academy events.

Effect of creating a long-lasting academy in cities (a) within 25 km from the hosting cities, excluding the hosting city, (b) between 25 km and 50 km from the hosting cities, excluding hosting cities and within the 0–25 km "donut," (c) between 50 km and 75 km from the hosting cities, excluding hosting cities and within the 25–50 km "donut"; estimated with Sun and Abraham (2021).
Note: Control group includes cities that never established a long-lasting academy and further away than 25 km, 50 km, and 75 km, respectively. Dependent variable is the logarithmic city population growth rate.



(A) (B) (C)

FIGURE 2.48: Test for spillover effects of big academy events.

Effect of creating a big academy in cities (a) within 25 km from the hosting cities, excluding the hosting city, (b) between 25 km and 50 km from the hosting cities, excluding hosting cities and within the 0–25 km "donut," (c) between 50 km and 75 km from the hosting cities, excluding hosting cities and within the 25–50 km "donut"; estimated with Sun and Abraham (2021).
Note: Control group includes cities that never established a big academy and further away than 25 km, 50 km, and 75 km, respectively. Dependent variable is the logarithmic city population growth rate.

Chapter 3

Flora, Cosmos, Salvatio: Pre-modern Academic Institutions and the Spread of Ideas*

While good ideas can emerge anywhere, it takes a community to develop and disseminate them. In premodern Europe (1084-1793), there were approximately 200 universities and 150 academies of sciences, which were home to thousands of scholars and created an extensive network of intellectual exchange. By reconstructing interpersonal connections that were made via institutional affiliations, we demonstrate how the European academic landscape facilitated the diffusion of ideas and led cities to develop: examples include botanic gardens, astronomical observatories, and Protestantism. Counterfactual simulations reveal that both universities and academies played crucial roles, with academies being particularly effective at connecting distant parts of the network. Moreover, we show that the diffusion of ideas through the network is remarkably resilient, even if we remove key regions such as France or the British Isles. In Europe, ideas gain prominence when they are channeled effectively by powerful institutions.

*This chapter is co-authored with David de la Croix (UCLouvain) and Rossana Scebba (UCLouvain/KULeuven). This article is published as a CEPR Discussion paper DP20569 at the following link: https://cepr.org/publications/dp20569.

3.1 Introduction

In Europe during the Middle Ages and Early Modern period, more than one hundred thousand scholars were engaged in the production, dissemination, and development of various forms of knowledge. These scholars did not operate in isolation: two key institutions, universities and academies, facilitated their interaction. These institutions organized teaching and research within self-governed communities of scholars, as stressed in Rashdall (1895) and McClellan (1985). From the halls of medieval universities such as the University of Paris and Bologna's *Alma mater* to, much later, lively debates in academies such as the Royal Society, formal institutions brought together scholars of diverse backgrounds to develop and disseminate scientific discovery. These institutions provided scholars not just with employment, but with a physical proximity that facilitated interaction.

In this paper, we measure how ideas spread through the academic affiliation network.[2] Our quantitative analysis models the diffusion of ideas through an evolving network of scholars with documented affiliations to formal institutions, based on data from the *Repertorium Eruditorum Totius Europae* project (RETE). Similar to the approach taken by Becker et al. (2024) in their analysis of the pre-WWII period, we base connections between scholars on their overlapping presence at the same academic institutions. To simulate the transmission of ideas, we combine an epidemiological approach with the network structure (Banerjee et al., 2013; Fogli & Veldkamp, 2021; Koher et al., 2016; Zamani et al., 2023). Our approach fits within the class of network diffusion models, such as that formalized in Bramoullé and Genicot (2024). Importantly, we take the network as given, focusing on the dynamics of diffusion rather than on the mechanisms of network formation.

Our network is dynamic and spans 1084 to 1793. Our timeframe starts with the establishment of Irnerius's school of jurisprudence in Bologna (c. 1050 - after 1125) and concludes with the French Revolutionary Convention (1793). The *nodes* of the network are premodern scholars. A connection, or an *edge* in the network, is established between any pair of scholars who share at least one year of concurrent affiliation at the same institution and work within a broadly similar field. We assume that ideas spread through networks like infections, with scholars transmitting their inventions to peers with a certain probability. In our model, idea diffusion is governed by a single parameter, the *link activation probability*. A low value suggests that face-to-face interactions rarely lead to knowledge transfer (e.g., due to non-academic discussions), while a high value indicates they are an effective channel for diffusion.

[2] See Borgatti and Halgin (2011) for a survey of papers using affiliation networks.

Ideally we would estimate this parameter by comparing the model's predictions to empirical moments that reflect the historical spread of ideas. However, a key challenge is that we do not directly observe ideas as they propagate. Unlike studies that trace the trajectory of a single idea—such as the study by Xue (2025) on Wang Yangming's influence in Chinese texts, or Giorcelli, Lacetera, and Marinoni (2022) on the spread of Darwinian theory via Google Books—or those that focus on specific linguistic or regional contexts, like Chiopris (2024) on 19th-century ideas in the German library consortium, our setting lacks detailed records of idea diffusion. Contemporary work, such as that of Ahmadpoor and Jones (2017), leverages citation and patent data to track knowledge flows over time: no comparable measures exist for Early Modern Europe. Still, we can observe outcomes plausibly linked to idea adoption, even though identifying consistent, city-level measures across Europe is especially challenging for the premodern period.

Our approach is to estimate the link activation probability by simulating the diffusion of ideas through the affiliation network and constructing *measures of exposure* at the scholar, institution, and city levels. We then correlate these simulated exposures with historical outcomes in what we term *auxiliary models*, drawing on the framework of indirect inference (see Smith (2008)). We chose two intellectual breakthroughs from the Scientific Revolution that have both historical significance and observable proxies, to benchmark our diffusion model.

The first auxiliary model is the rise of botany as an independent discipline, driven by Leonhart Fuchs (1501–1566), a professor at the Universities of Tübingen and Ingolstadt. Fuchs' work emphasized direct observation of nature, culminating in the publication of a comprehensive herbal featuring accurate plant illustrations and medicinal descriptions. We label this shift "Botanical Realism". The heightened interest in botany across Europe that followed is evidenced by the spread of botanic gardens. We calculate each city's simulated exposure to Botanical Realism and use a proportional hazard model to estimate the probability of a botanic garden being established. We find that this probability increases with exposure, which is consistent with the model being accurate.

The second auxiliary model focuses on the astronomical revolution, particularly the foundational role of Johannes Regiomontanus (1436–1476) in the advancement of trigonometry and astronomy. We group his innovations, which were later instrumental to the work of Copernicus, Kepler, and Galileo, under the label "Mathematical Astronomy". Regiomontanus held positions in Vienna, Bratislava, Padua, and Rome, which facilitated the transmission of his ideas. We measure exposure to his innovations and analyze the correlation with the *creation of astronomical observatories* across Europe. Again using a proportional hazard model, we find a positive relationship between exposure and the likelihood of an observatory being established.

In both cases, we control for Euclidean distance from the place of origin of the idea, accounting for diffusion via geographic proximity (consistent with gravity models), which would be the case, for example, for books (Dittmar, 2011). Our exposure measures remain predictive even after this control, which highlights the critical role of social and institutional connections in the transmission of ideas. This finding echoes apprenticeship models, in which the master-apprentice face-to-face communication is key (De la Croix, Doepke, & Mokyr, 2018), but also contemporary insights, such as those of Atkin, Chen, and Popov (2022), on the persistent importance of face-to-face interactions in environments like Silicon Valley.

We then estimate the link activation probability by maximizing the joint likelihood of these two auxiliary models. Next, we use the model to carry over additional empirical analysis, going beyond the Scientific Revolution and the direct relation between ideas and outcomes, to cover an example of a backlash against an idea. Scholasticism was pioneered by Petrus Lombardus (c. 1100 – 1160), a professor in Paris, and it was the dominant approach to philosophy and theology in the Middle Ages. Followers of scholasticism used logical reasoning to explore theological questions, and this method was adopted in many universities. Over time, however, it became increasingly detached from the practical concerns of believers and devolved into abstract debates, a decline often cited in historical literature (Barrett, 2023; Chaunu, 2014). It is possible that Protestantism emerged as a reaction to Scholasticism, emphasizing the importance of scripture over intellectualized theological debate (Chaunu, 2014). To test this hypothesis, we simulate the diffusion of Scholasticism through the affiliation network and measures of exposure. We use 1508, just before Martin Luther's academic career began, as the start of the simulation to minimize potential confounders. We capture the exposure to Scholasticism across universities in 1508 and infer the exposure of nearby cities. Using a linear probability model, we assess whether cities with higher exposure to Scholasticism were more likely to embrace Protestantism using the dataset and the control variables from Rubin (2014). The results show a strong correlation between exposure to Scholasticism and the likelihood of becoming Protestant.

In a similar vein, we extend the classical finding that pogroms against Jews are more likely to occur following major shocks—such as the Black Death (Becker & Pascali, 2019; Jedwab, Johnson, & Koyama, 2019; Voigtländer & Voth, 2012)—by showing that the anti-Judaic ideas embedded in scholastic thought acted as a complementary force in this process. In addition, we examine a distinct case involving a demonstrably false belief: the claim that Swedes are descendants of the lost civilization of Atlantis.

Having demonstrated the model's effectiveness by comparing its predictions with real-world outcomes, we turn to exploring counterfactual scenarios. In this analysis, we compare observed outcomes with the hypothetical scenarios that would have emerged under alternative conditions. These conditions include:

assigning the invention of a given idea to a different scholar within the network, removing affiliations to academies or to institutions situated in a specific geographical area, and excluding scholars belonging to the Jesuit community. Our aim in examining these counterfactuals is to assess the network features that are most critical for idea diffusion.

In the first experiment, we reassign the intellectual origin of an idea to different individuals and ask whether it would still propagate across Europe. While some peripheral institutions are not well enough connected to guarantee the survival of an idea, we find that in most cases, ideas still reach the entire network within a couple of centuries. However, the speed and route of diffusion vary, emphasizing the non-ergodic nature of the process and the importance of initial conditions. In the second experiment we compare the role of academies to that of universities. When we remove academies from the network, there is a notable drop in the geographical reach of ideas, especially those that originated in remote or disrupted areas. Academies, with their international memberships, often served as critical diffusion channels, bridging regions and sustaining intellectual exchange during crises such as university closures during the Thirty Years' War. In the third experiment, we simulate the removal of entire countries or regions—such as the British Isles, France, or the Italian and Iberian Peninsulas—to assess their systemic importance. Surprisingly, even when major regions are excluded, most ideas continue to spread widely across Europe. This suggests that most regions were not essential to the flow of knowledge, underscoring the network's remarkable resilience to local shocks.

Beyond the influence of wars and academies, these simulations highlight the serendipitous nature of knowledge transmission. At times, an idea follows a remarkably narrow path, where sheer luck determines its survival.

It is important to emphasize that our focus on counterfactual scenarios is primarily methodological. We use counterfactual scenarios to gain a deeper understanding of the academic network's intrinsic properties, rather than to definitively predict what might have occurred in these hypothetical situations. This approach aligns with the spirit of classic works like Fogel's (1964) seminal study on the impact of railroads on American economic development, which emphasized the importance of counterfactual reasoning for historical analysis.

Before delving further into the specifics of our approach, it is important to acknowledge that institutional overlap, while insightful, captures only a fragment of the multifaceted reality of knowledge dissemination in the premodern period.[3] We primarily focus on scholarly interactions within formal institutional settings during the medieval and premodern periods. These exchanges occurred through

[3]Aggregation centers were crucial for intellectual exchange in the medieval and premodern world. As shown by Brunt and García-Peñalosa (2022), cities played a similar role in fostering knowledge diffusion by concentrating individuals and increasing opportunities for encounters and idea transmission.

in-person interactions as well as institutionally-driven epistolary communication, as practiced in some academies. We emphasize the structured environment of these institutions, which facilitated scholarly interaction and fostered intellectual exchange. For instance, academy correspondence was directed to all members through official channels rather than through private, intentional communication. Beyond formal communication, these institutions also provided spaces for direct intellectual engagement among individuals.[4]

While our model prioritizes scholar-to-scholar interactions within institutions, we recognize that, in the realm of in-person intellectual exchanges, student-teacher relationships, as well as student associations (such as *nationes*, *bursae*, fraternities and others) also played a part in the dissemination of ideas (see Koschnick (2025) on teacher-student interactions in Oxford and Cambridge). However, reconstructing comprehensive student attendance records from the medieval and Early Modern periods is a highly complex task, not only due to incomplete or fragmentary data but also because students' participation in university communities was more transitory than that of professors.

Other avenues, such as reading habits, undoubtedly contributed to knowledge dissemination, but quantifying their impact remains challenging. Citations could offer insight into what was being read and discussed in the academic community (Zhao & Strotmann, 2015), but they would not provide a complete picture of reading patterns.[5] While citations can reflect "positive" engagement with prior work, including critical remarks, they do not capture instances where scholars deliberately avoided acknowledging influential ideas. Not all works that were read were cited, and this underscores the limitations of citation analysis for reconstructing the full spectrum of intellectual interaction. Our study, therefore, focuses on traceable, institutionally-mediated pathways through which ideas spread. Unlike in biological contagion, 'infection' as an analogy for idea diffusion does not imply endorsement or adoption. As Banerjee et al. (2013) highlight in their study on microfinance diffusion, individuals may receive or transmit information without endorsing it. For this reason, we prefer to speak of exposure rather than infection.

This distinction also informs our methodological approach. In the spirit of an intent-to-treat (ITT) framework, we focus not on actual compliance with ideas, but on the potential for exposure. By observing whether scholars were present in the same location at the same time—in universities or academies, which typically had a limited number of scholars—we can infer exposure to ideas,

[4]A striking example of this dynamic is the relationship between the abovementioned Regiomontanus and Polish astronomer Martinus Bylica de Ilkusz, who met at the University of Padua in 1463. Their long-lasting intellectual bond—with Bylica amending the Regiomontanus manuscripts—mirrored that of their mentors, respectively, Georg Peuerbach and Martinus Król (Domonkos, 1968), and underscores how institutional settings also fostered scholarly connections across generations.

[5]Moreover, in premodern texts, citations did not follow standardized formats, making it difficult to systematically trace citations across different works.

regardless of whether these ideas were ultimately adopted (i.e., compliance). This approach contrasts with much of the existing literature on the diffusion of knowledge, which relies on measures such as the actual content of publications, correspondence, citations, translations, or co-authorship, all of which inherently assume compliance (Abramitzky & Sin, 2014; Donker, 2024; Goyal, Van Der Leij, & Moraga-González, 2006; Hallmann, Hanlon, & Rosenberger, 2022; Roller, 2023). In randomized controlled trials, ITT is generally preferred for primary analysis because it avoids the biases that arise from excluding non-compliant participants and preserves the benefits of randomization. In ITT analysis, the estimate of treatment effect is generally conservative (Gupta, 2011). We apply this same principle to trace the diffusion of ideas.

Another advantage of our approach is that gaps in the affiliation network are identifiable. Gaps occur when universities are underrepresented in our sources, causing some professors to be missing. By contrast, gaps in correspondence networks—such as letters lost to history—are inherently unknown and thus impossible to assess.

In addition to the literature on network structures discussed above, our paper connects to several other strands of research. First, it engages with the history of science, from broad studies on the emergence of science and the scientific method (Needham, 1964; Wootton, 2015), to analyses of scientists' roles in society (Ben-David, 1971; Hanlon, 2025), and to work on shocks to the market for ideas and technology. Mokyr (2005, 2011b, 2016), Ó Gráda (2016), and Almelhem et al. (2023) explore the roots of the industrial revolutions through the accumulation and application of useful knowledge during the Scientific Revolution and Enlightenment. Second, we contribute to the literature on the economics of innovation. Historical scholarship highlights the impact of highly skilled individuals (Meisenzahl & Mokyr, 2012), specialized engineers (Hanlon, 2025; Maloney & Valencia Caicedo, 2022), and key inventors (Hallmann, Hanlon, & Rosenberger, 2022) in driving Britain's technological edge. Other works show how the institutional context matters: classical composers thrived in more liberal environments (Borowiecki, 2013), patent systems shaped the direction of innovation (Moser, 2005), and external shocks—such as the U.S. Civil War—spurred targeted technological responses in Britain's textile industry (Hanlon, 2015). Our approach aligns with Akcigit et al. (2018) in its focus on tracking "individuals, their productivity, and their interactions over time" (Akcigit et al., 2018, p.2). Our analysis also relates to work on academic superstars, which shows that the removal of central figures can disrupt or reconfigure knowledge networks (Azoulay, Fons-Rosen, & Zivin, 2019; Azoulay, Graff Zivin, & Wang, 2010). While the focus of those papers is on publication output, the underlying question—how resistant to disruption academic networks are—is also relevant to our framework.

Having described the strength of European academia in terms of connectivity and resilience, we believe it could have been instrumental to Europe's success during the Early Modern period. This remains speculative, however, as we lack comparable data on academic networks in other parts of the world, and rely only on anecdotal evidence. In 1798, Thomas Malthus (1766–1834), a fellow at the University of Cambridge from 1793, published a treatise on population and development (Malthus, 1807). He developed the idea that population growth tends to outpace food production, leading to inevitable constraints to development. In 1818, Malthus became a Fellow of the Royal Society. Malthus' view had an immense influence on political economy in the following decades. Still today his ideas are modeled and debated (André & Platteau, 1998; Ashraf & Galor, 2011). At about the same time, Hung Liang-Chi (1744–1809), a high official of the Chinese imperial administration, developed similar ideas. The ideas were particularly relevant for understanding Chinese dynamics in the 19th century. Still, Hung Liang-Chi largely disappeared from the record and only rediscovered in the 20th century (Silberman, 1960). How can we understand the differences in the fates of these two ideas? The approach we develop in this paper can be applied to explain Malthus' success. Malthus was integrated into the broad European academic network, where his ideas could spread. Hung Liang-Chi belonged to an administration in which ideas were developed by individuals but not subject to broad dissemination and discussion.

The paper is organized as follows. In Section 3.2 we present the methodology, including the construction of the database and how it is mapped into an affiliation network. We also present the epidemiological model to simulate the diffusion of ideas. In Section 3.3 we detail the structural estimation of the model parameter, based on *flora* and *cosmos*. In Section 3.4 we present further empirical assessments, based on *salvatio* (Scholasticism). The counterfactual experiments are detailed in Section 3.5. Section 3.6 concludes.

3.2 Data and Methodology

We now present our methodology, starting with the compilation of the database of professors, followed by the definition of the temporal network and the epidemiological model that we use to describe how ideas flow.

3.2.1 83,000 scholars

Our comprehensive database of scholars comprises information on 83,000 individuals spanning the period 1000-1800. The data were collected manually from 669 distinct sources. Unlike other studies that rely on ex-post recognition of scholars—such as that derived from Wikipedia/Wikidata (see Laouenan et al.

(2022) and Serafinelli and Tabellini (2022))—our selection is based on membership lists or secondary sources related to key higher education institutions. These institutions fall into three categories: universities (referenced in Frijhoff (1996); see also De la Croix et al. (2024)), scientific academies (as cataloged in McClellan (1985), and further discussed in Zanardello (2024)), and various other institutions with links to universities, including Italian Renaissance academies mentioned in The British Library (2021), and other higher education entities that conferred academic degrees.

Medieval universities primarily focused on four disciplines: theology, law, arts and humanities, and medicine. The faculty of arts provided foundational education to grammar school pupils, many of whom became teachers themselves, contributing to rising literacy rates among the general population. Some students progressed to higher faculties, preparing for professions in other fields. The faculty of medicine trained medical practitioners, the faculty of laws produced future administrators with specialized knowledge in canon or civil law, and the faculty of theology trained teachers for episcopal schools, where ordinary parish priests were instructed (Pedersen, 1992). Academies, emerging later in the 17th and 18th centuries, were created to foster new areas of research not traditionally covered by universities (Applebaum, 2003; McClellan, 1985). These ranged from informal groups of amateur naturalists or local historians, to prominent official societies that gathered leading scholars, published journals, and formed networks of corresponding members, known collectively as the Republic of Letters (Mokyr, 2016).

To compile the list of scholars from each academy and university, we mostly relied on secondary sources, mainly books on the history of these institutions and their members, which were themselves based on primary records. For universities, our aim was to include scholars involved in teaching, covering a range of positions from royal chairs in France to fellowships in England. Further details on the inclusion criteria for university scholars can be found in De la Croix et al. (2024), while global statistics are available in De la Croix (2021) and various issues of the *Repertorium Eruditorum Totius Europae*. For academies this process was generally straightforward, since comprehensive membership lists are often available. Our data on academies have already been utilized in works such as Blasutto and De la Croix (2023) for Italian academies, De la Croix and Goñi (2024) for analyzing father-son pairs across academies and universities, and Zanardello (2024) for evaluating the impact of different fields of study within academies. These lists encompass several membership categories, including ordinary, corresponding, and honorary members. Corresponding members, though not present at academy meetings, contributed from a distance. Honorary members often included local dignitaries like bishops, wealthy merchants, and governors, who supported and protected the academies. To prevent skewing our results due to the inclusion of these sometimes prominent figures, we excluded anyone holding honorary membership or those who were clearly not scholars or intellectuals (e.g., Napoleon, who was elected to the Académie des Sciences in 1797).[6]

The resulting database can be accessed at https://shiny-lidam.sipr.ucl.ac.be/scholars/.

Additionally, we leverage data from VIAF and Wikipedia entries associated with each person in the database, when available, to provide a unique measure for individual quality (Curtis & De la Croix, 2023).[7] This variable is intended to reflect a scholar's influence and recognition, and is obtained using Principal Component Analysis to create a single "human capital index" score for each individual. Later in the paper, we will refer to this measure and employ it in our analysis.

3.2.2 Definition of the affiliation network $\mathbb{G}$

We now look at the data on scholars and their affiliation to institutions with reference to network theory, a powerful tool for studying the spread of information over time and space, among other subjects (Goyal, 2023; Jackson, 2008). We model the affiliation network as a graph, where *nodes* represent scholars and *edges* denote their contemporaneous presence at the same institution. This network is derived from an initial bipartite representation, consisting of two types of nodes: scholars and institutions. In the bipartite version, edges connect scholars to the institutions they were affiliated with. Since our focus is on scholar-to-scholar interactions, we project this bipartite graph onto a single-mode network of scholars, where an edge between two nodes represents their concurrent affiliation with the same institution, as shown in Figure 3.1. Premodern universities were small, and it is reasonable to assume that all the professors knew each other. Appendix 3.7 presents some descriptive statistics for the 10 institutions with the highest number of scholars in 1793, differentiating by type (university vs academy). The average university has fewer than 60 professors. Academies hosted a higher number of scholars on average, reaching 80 in Halle and 122 in London, but the average number of scholars at academies is much more variable, reflecting a more flexible structure and a higher number of corresponding members.

Our analysis spans the foundation of a new school of jurisprudence, which would later become the University of Bologna, in $\underline{t} = 1084$, to the French

[6]Members were admitted through an election process, which varied slightly between academies in terms of the required quorum. Typically, the process began with a nomination by existing members, which was recorded in the academy's minutes along with the scheduled election date. On that day, the academy's president ensured the quorum was met, after which the members voted (Applebaum, 2003; Gunther, 1925).

[7]VIAF (Virtual International Authority File) is a database that connects names of people, organizations, and titles from library catalogs around the world into one shared system.

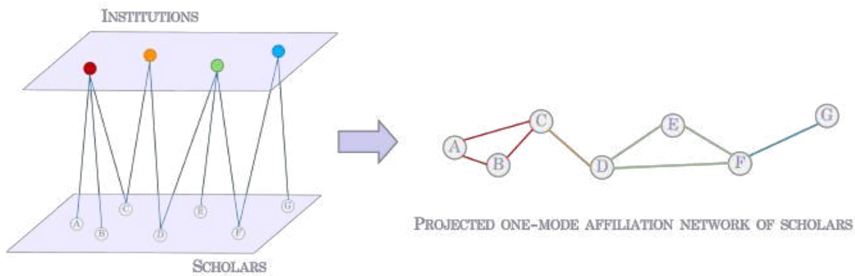


FIGURE 3.1: Intuitive representation of network projection: from a bipartite or two-mode graph to the one-mode affiliation network of scholars. Diagram inspired by (Geraerts & Vasques Filho, 2024).

Convention in $\bar{t} = 1793$, which led to the abrupt closure of all universities and academies on the territory of the new Republic. During this timeframe, the network's nodes (scholars) and edges (connections) existed only within specific periods defined by the duration of each scholar's activity and their affiliations with institutions.

More formally, given two scholars i_s and i_v, the link between i_s and i_v lasts as long as i_s and i_v share an overlapping period of affiliation at the same institution. This implies that the collection of edges is dynamic over time: edges serve as channels for the spread of ideas, appearing and disappearing only while scholars are active at the same institution. In contrast, nodes (scholars) exist in the network as long as they are active, i.e. affiliated with one or more institutions as in our main database. Figure 3.2 shows the evolution of the number of active scholars over time, showing overall exponential growth, in particular after 1650, with the emergence of academies.

The active period of each scholar commences at the start of their academic career and concludes at their retirement or death. Activity begins in the earliest known year of affiliation with a formal educational institution, when available. For university professors, this is the year they began teaching, while for academy members, it is the year they were elected as members of the academy. If the exact affiliation date is unavailable, we infer the first year of affiliation from approximate dates. In more extreme cases, we use the earliest available date among: 30 years after the birth year, the year of death, the institution's closing date, or 1793, which marks the end of our study period. This approach aims to provide a conservative estimate of each scholar's active period.

Scholars cease to be active when they leave the institution, if this date is available. For university professors it is the year their teaching ends. For members of academies it is usually the death year, or in some rare case, the

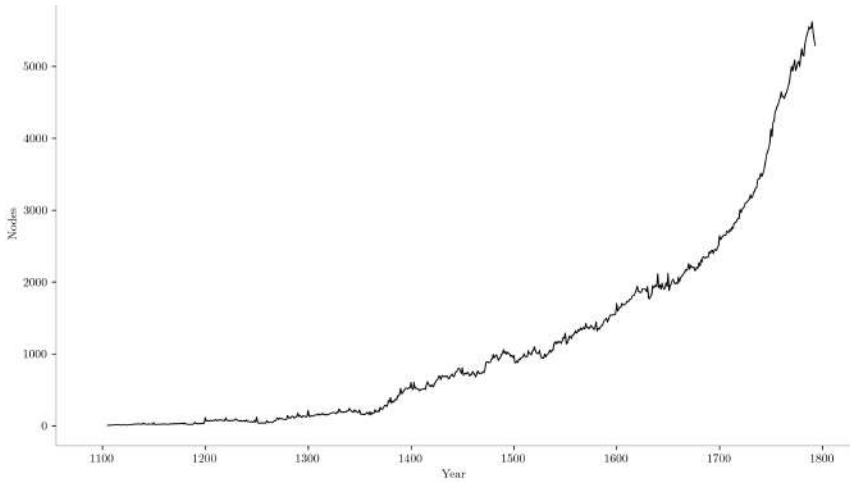


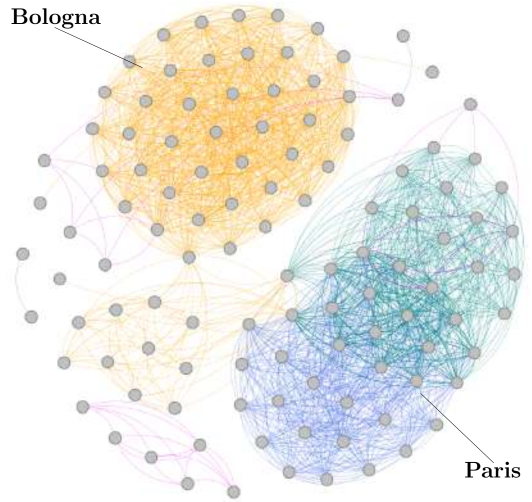
FIGURE 3.2: Number of active scholars in the network, 1084-1793.

year the member was expelled. When there is no precise information about the end of their activity, we infer it in one of two ways for university professors and academicians. For university professors without a precise end affiliation date, we assume it is equal to the approximate affiliation date when available. Otherwise, if these pieces of information are unknown, we assume that a university professor will teach in that university for eight years.[8] Hence, we take the earliest date between the beginning date (after adding eight years) and the year of death. For academicians without a precise end date of their affiliations, we also assume it from more approximate dates if available. When not possible, for these scholars, we assume a lifelong affiliation,[9] so we take the year of death when there is no more precise information. Otherwise, when we do not know the year of death either, we assume they stay in the academy only one year, imposing the end affiliation date equal to the beginning date.

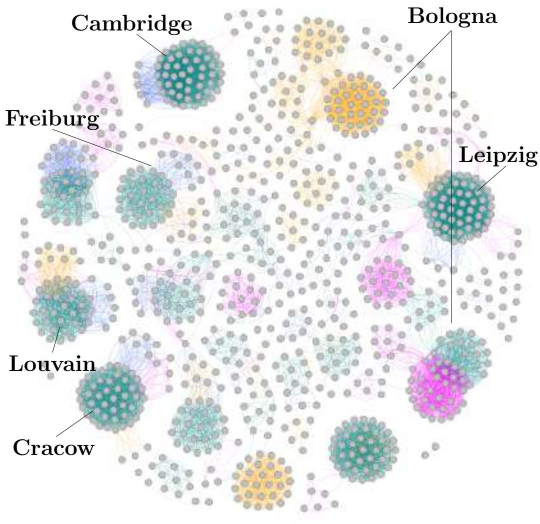
The affiliation network reveals several important characteristics, visualized in Figure **??**. Before the rise of academies, the network at any given time typically consisted of a series of mostly disconnected clusters, each representing a single university. These clusters in turn comprised *cliques* (fully connected clusters) of scholars operating in the same field at the same university. Occasionally,

[8]Eight years being the median of the affiliation years in the sample of university professors for whom we know the precise beginning and end affiliation dates. This data is consistent with the literature: Koschnick (2025) finds that the median length of academic careers at Oxford and Cambridge is 9 years.

[9]Academies usually grant a lifelong affiliation.

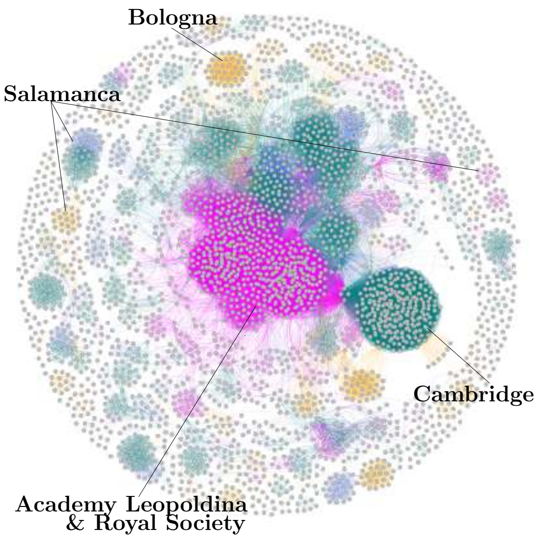


(A) 1200

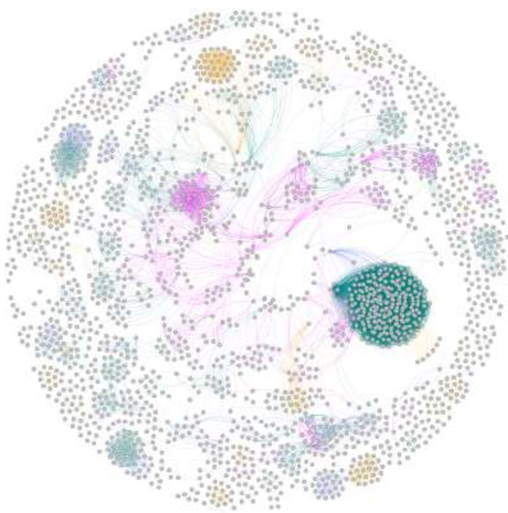


(B) 1508

FIGURE 3.3: Affiliation network snapshots in 1200 and 1508. Edge colors: theology (blue), law (orange), humanities (teal), sciences (magenta). Isolated nodes not shown.



(A) 1730



(B) 1730 (no academies)

FIGURE 3.4: Affiliation network in 1730 with and without academies. Edge colors: theology (blue), law (orange), humanities (teal), sciences (magenta). Isolated nodes not shown.

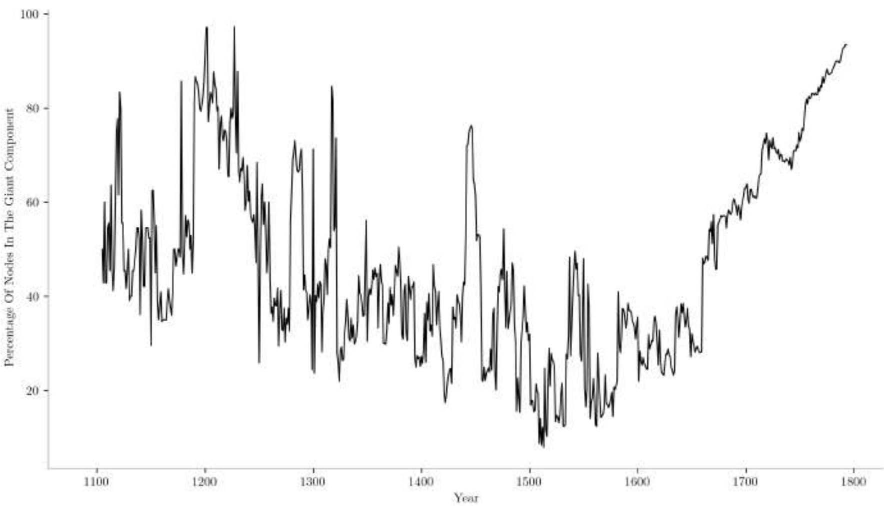


FIGURE 3.5: Percentage of active scholars in the giant or connected component over time, 1084-1793.

cliques overlapped, highlighting scholars active in multiple fields. Over time, as scholar mobility increased, occasional links began to form among different university clusters, creating pathways between otherwise separate regions of the network. With the emergence of academies, however, connections between scholars multiplied significantly. Academies often appointed foreign members, serving as bridges between previously isolated universities clusters: Figure 3.4b depicts the 1730 network in a scenario where academies are removed. From the 1650s to the end of the timeframe, the cluster- and clique-based appearance of the early network transformed into a densely interconnected web, and this is attributable to the academies.

This effect is also visible in the share of nodes connected to the main network, known as the giant component—the largest connected subgraph in which any two nodes are mutually reachable—as shown in Figure 3.5. In the early period, around 1200, the affiliation network was nearly fully connected, reflecting a small number of institutions and high scholar mobility. From the 14th century onward, network connectedness followed a general downward trend, punctuated by occasional spikes. These spikes typically occurred when a mobile scholar bridged otherwise disconnected parts of the network. The size of the giant component reached a low point during the Reformation, which curtailed mobility across religious lines, as discussed in De la Croix and Morault (2025). Although the Thirty Years' War (1618–1648) did not drive a further decline, it prolonged the fragmentation, with the giant component remaining relatively

TABLE 3.1: Network statistics for the affiliation network, 1200-1700.

Year	1200	1300	1400	1500	1600	1700
Total scholars	114	209	602	976	1702	2644
Avg. degree	30.54	15.86	35.27	25.10	35.45	48.30
Std. dev degree	14.57	9.43	30.82	17.50	62.93	71.03
Giant size	104	149	164	312	607	1669
Giant %	91.2	71.3	27.2	32	35.7	63.1
Second-largest comp	6	22	113	138	212	51
Clustering coef	0.9	0.91	0.92	0.91	0.88	0.89
Avg. distance (giant)	2.16	3.68	1.95	4.02	5.86	4.25
Std. dev distance	0.99	1.81	0.95	1.86	3.18	1.84

small. Connectivity recovered with the rise of academies, as the share of nodes in the giant component increased from 40% in 1650 to 90% by 1793.

In Table 3.1, we report network statistics at the beginning of each century, between 1200 and 1700, following Goyal, Van Der Leij, and Moraga-González (2006), to assess whether our network exhibits properties of an emerging small world. The graph consistently displays a high clustering coefficient (above 0.88 in all periods), well above the levels expected in a random graph with similar size and density, where clustering would typically approximate to the average degree divided by the number of nodes–and hence close to zero in sparse networks of this scale. At the same time, average path lengths (or distances) within the giant component remain low–between 2.16 and 5.86–and stable over time, even as the number of scholars increases significantly. While the network is highly fragmented in earlier centuries, with the giant component capturing as little as 27.2% of scholars in 1400, its expansion to over 60% by 1700 reflects a growing integration of academic communities across institutions. In other words, in the earlier centuries, the affiliation network resembles an archipelago of small "islands" (i.e. universities) which are internally well connected (as shown by the very high clustering coefficients) but connected to each other only via the occasional mobility of scholars. Over time, particularly by 1700, these isolated islands began to form bridges–thanks to multiple simultaneous affiliations, which was due to the emergence of academies. The result is an increasingly interconnected network, where the size of the giant component grows to nearly two-thirds of all scholars, while clustering remains high and path lengths stay short. The size of the second-largest component further underlines this aspect: it increases up to 1600 but drastically decreases in 1700, thanks also to the arrival of scientific academies. This shift marks the emergence of a small-world structure out of what was once a fragmented archipelago.

This affiliation network is the structure on which we will simulate the spread of ideas. Each idea is assumed to originate in a specific year and from a particular inventor, who can transmit it to their neighbors at each time step, but under certain conditions. Each idea belongs to a broad field that reflects the main disciplines of premodern times, and can spread only among scholars active within the that field. Our assumption is that if a scholar is working in science[10] they are presumed to engage productively with peers in medicine or applied science.[11] One may argue that in the past there were all-round scholars who were equally fluent, for example, in both science and philosophy. Still, we decide to assign specific fields to each idea, while acknowledging that this may cause us to underestimate the speed of diffusion of ideas.

3.2.3 Ideas, inventors and exposed scholars

In this paper, we simulate the spread of ideas originating from inventors, i.e., scholars who developed a new idea at a specific point in time. Inventors can propagate their ideas to their peers, who, once exposed, can further propagate them to others. Upon being exposed to the idea, scholars can pass it along to their own neighbors without needing to maintain an enduring direct link with the original inventor. In our context, an inventor is a scholar recognized for a groundbreaking idea, as reported in one of the major historical encyclopedias. We use English (2005) for ideas spread before 1500 C.E., and Applebaum (2003) for ideas diffused between the invention of the printing press (circa 1450s) and the French Revolution. From these sources, we identify some significant ideas that changed the course of history, prioritizing those for which historical outcomes are available to validate our model's predictions. For each idea we pinpoint the inventor, as detailed in Section 3.3.2, Section 3.3.3, and Section 3.4.1.

We simulate the spread of three main ideas—*flora*, *cosmos*, and *salvatio*—two from the Scientific Revolution and one from the Middle Ages, and compare simulated exposure to observed outcomes. For the Scientific Revolution, we draw on Applebaum (2003), focusing on the earliest ideas for which European-level outcome data is available. In astronomy, the first idea is attributed to Regiomontanus, whose *Mathematical Astronomy* (developed from 1450, published in 1496) formalized Ptolemaic models for future research. Its corresponding outcome is the founding of astronomical observatories. In botany, Applebaum (2003) identifies *Botanical Realism*, foundational to Fuchs's 1542 herbarium; we link exposure to this idea with the creation of botanic gardens.

[10] Science includes mathematics, logic, physics, chemistry, biology, astronomy, earth sciences, geography, and botany.

[11] Applied science includes engineering, architecture, and agronomy.

For the Middle Ages, we identified key academic ideas using the index in English (2005). These include alchemy, anatomy (including practical surgery), astrology, computus, civil law, economic thought, cartography, humanism, music, optics, political theory, punctuation, and the Scholastic method. To balance the two scientific ideas above, we select one from a different domain: theology. *Scholasticism*, rooted in the work of Lombardus, emphasized rigorous logic and dialectical reasoning to reconcile faith with reason. It fostered a systematic approach to inquiry that shaped both scientific and philosophical thought. At the same time, it spurred theological backlash, especially from movements like Protestantism that emphasized scriptural authority over rational deduction. The associated outcome in our empirical analysis is the probability of a city becoming Protestant.

A key challenge, common to all three ideas, is determining when the scholar first developed the concept. To define this moment, we try to identify two dates for each idea: (i) the publication date, which refers to when the scholar first published a work on the topic, and (ii) the inception date, which is the year when the scholar first conceived the idea and likely began discussing it with colleagues. We identify the inception date manually, by reviewing biographical information and related historical context.

In our model ideas spread via interactions among scholars, and these interactions can only occur when scholars are alive. Therefore, using the inception date rather than the publication date allows us to better capture the dynamics of idea dissemination through the scholarly network, as it reflects the period when discussions and exchanges of the idea were possible. Our preferred date is therefore the inception date, though we rely on the publication date when the inception date is unavailable.

Overall, ideas are more likely to spread if certain conditions are met: (a) scholars have long lifespans, giving them more time to spread their ideas; (b) there is a high density of scholars at a given institution, creating more opportunities for intellectual exchange, and (c) scholars move between institutions, which facilitates the dissemination of ideas across different scholarly communities. However, the spread of each idea may vary significantly depending on specific factors, including the centrality of the inventor within their peer network, their affiliations with large institutions, and the timing of the idea's inception.

To briefly clarify some terminology: we use inventor as shorthand for the main proponent or originator of an idea—not necessarily an inventor in the traditional sense, but often someone who developed, articulated, or popularized a concept. Similarly, idea is used broadly to encompass various intellectual contributions, including theses, paradigms, and methodological approaches, which differ in scope, complexity, and impact.

3.2.4 Epidemiological model

Following the view that social networks diffuse information like infectious diseases (Banerjee et al., 2013; Fogli & Veldkamp, 2021), we start from an epidemiological approach. There is a fixed number of nodes, N, each representing a scholar. Time is discrete, with $t \in \{\underline{t}, \ldots, \bar{t}\}$, $\underline{t}$ and $\bar{t}$ being the start and end dates of our analysis. At each date, a node can be susceptible or infectious.[12] A contact between two nodes appears as a undirected link in the network at a given time. Interactions are represented by an adjacency matrix $A_t = [a_{sv}]_t$ of dimension $N \times N$, with each element a_{sv} taking value 1 if scholar s and v are connected at time t, and zero otherwise. Connections will depend on whether s and v are working in the same field at the same time in the same institution (more on this later). We represent a temporal network $\mathbb{G}$ by a set of adjacency matrices A_t.

The state of the world is described at each date by a vector $I_t = [i_s]_t$ of length N. We only have binary entries in I_t, with $i_s = 1$ if scholar s is infected, and $i_s = 0$ otherwise. Initially, there is no idea and nobody is infected. At some date t_0 an initial "inventor" has an idea. We thus have $[i_s]_t = 0$ for all $t < t_0$, and $[i_{s^\star}]_{t_0} = 1$, where $s^\star$ is the inventor.

Following the binary nature of the state vector, we use Boolean arithmetic, i.e. element-wise addition and scalar multiplication are replaced by the logical "or" and "and", respectively (Koher et al., 2016). Dynamics are then represented by:

$$I_{t+1} = A_t I_t + I_t \tag{3.1}$$

To understand this formula, consider the s scholar. If they are alive at period t, their infection status at $t+1$ is given by $\sum_v a_{sv} i_v$. With Boolean arithmetics, this term is equal to 1 if there is at least one v such that $a_{sv} = 1$ (s has met v) and $i_v = 1$ (v is infected). If, instead, s is either unborn or dead at time t, $a_{sv} = 0$ $\forall v$, and their infection status does not change.[13]

We also assume that once contaminated by an idea, a scholar cannot forget it. Hence the "recovered" state of the epidemiological model is not relevant here.

So far we have assumed that ideas are transmitted upon contact with probability 1. If instead, there is a link activation probability $\alpha \in [0, 1]$,[14] we define a

[12]Here, our model closely relates to Banerjee et al. (2013), since infection does not equate adoption of the idea.

[13]While the model could in principle track how many times s has been exposed to infected neighbors, as suggested by Bramoullé and Genicot (2024), we adopt a simplified binary-state process: infection occurs upon the first effective contact. Subsequent contacts with infected peers do not accumulate and have no further effect on the "intensity" of infection.

[14]Rather than assuming automatic transmission, we model a probabilistic approach to idea diffusion, reflecting the uncertainty and selectivity observed in historical intellectual exchanges–a logic similar to the stochastic imitation dynamics in Brunt and García-Peñalosa (2022).

stochastic operator $\Omega^d(A)$ (following Koher et al. (2016)) which acts element-wise on the adjacency matrix: for $a_{sv} = 0$, we have $\Omega^d(a_{sv}) = 0$; for $a_{sv} = 1$, we have $\Omega^d(a_{sv}) = 1$ with probability α and $\Omega^d(a_{sv}) = 0$ with probability $1 - \alpha$. Each potential transmission is evaluated independently on each edge: a susceptible scholar becomes infected through contact with an infected neighbor with probability α, based on an independent draw.

Dynamics of the state vector I are now represented by:

$$I_{t+1}^d = \Omega^d(A_t)I_t^d + I_t^d \tag{3.2}$$

where d is an index of simulations (draws). Since each 1 in A_t independently survives with probability α, the expected value of the stochastic contact matrices is:

$$\mathbb{E}[\Omega^d(A_t)] = \alpha A_t.$$

Such a specification increases the computational effort and allows for interactions between topological effects (those coming from the structure of the network) and probabilistic effects.

We now define three different levels of exposure. These levels are expected levels, given the stochastic nature of the simulations.

Expected scholar s exposure $[\bar{i}_s]_t \in [0,1]$ is obtained by averaging individual exposure over D simulations:

$$[\bar{i}_s]_t = \frac{1}{D}\sum_{d=1}^{D}[i_s^d]_t$$

Expected institution k exposure $S_t^k \geq 0$ is obtained as an average over individuals s belonging to set of members $V(k,t)$, at time t, weighting individual exposure by quality q_s:

$$S_t^k = \sum_s \underbrace{q_s}_{\text{quality}} \left(\underbrace{I(s \in V(k,t))}_{\text{membership}} \; \underbrace{[\bar{i}_s]_{t'}}_{\text{exposure}} \right) \tag{3.3}$$

The quality variable q_s is derived from footprints left in the libraries, as described above in Section 3.2.1. It is a comprehensive measure of human capital (see De la Croix et al. (2024) and Curtis and De la Croix (2023) for more details), as reflected in lifetime achievements.

Accounting for institutional exposure being influenced by the publication output of scholars implies that better scholars, with higher quality, contribute more to the institution's exposure compared to scholars with lower q_s. Consequently, if a scholar did not publish anything over their lifetime, resulting in a zero quality index, they will not contribute to the institution's exposure.

The measure of exposure S_t^k will be used in the proportional hazard models of Sections 3.3.2 and 3.3.3 to assess how exposure at a certain date is correlated with the emergence of botanic gardens or observatories.

It is also useful to define an exposure measure at the institution level which takes into account a window of time (instead of a point in time), which will be used in Section 3.4.1 and Appendix 3.C.3. We opted for a window of 30 years—one academic generation: the average age at appointment for university professors is 31 years, and their average age at death is 63; meanwhile, academicians begin their careers at an average age of 38 and typically pass away at 67 (Zanardello, 2024). Past this window an idea could survive the passing of its author, for example by persisting in the teaching material and/or as an influence on the culture of the institution. Accordingly, we define

$$\tilde{S}_t^k = \sum_s \underbrace{q_s}_{\text{quality}} \left(\frac{1}{30} \sum_{t'=t-30}^{t} \underbrace{I(s \in V(k,t'))}_{\text{membership}} \underbrace{[\bar{i}_s]_{t'}}_{\text{exposure}} \right). \tag{3.4}$$

Finally, **Expected city c exposure** $S_t^c \geq 0$ is obtained by averaging over nearby institutions, weighting by inverse distance:

$$S_t^c = \sum_k w_{ck} \tilde{S}_t^k \tag{3.5}$$

The weights w_{ck} are derived from the inverse distance between all the institutions in our database and the cities in our samples (precise details about these cities data are provided in each experiment). Considering the inverse distance means that the further a city is from an exposed institution, the lower the influence that reaches the urban center. An institution fully influences cities within 10 kilometers: essentially the city hosting that institution. Beyond 10 kilometers, the influence power decreases linearly, up to 1000 kilometers. After this threshold, we assume that the institution's influence loses all its power, reaching a weight of zero, and thus it cannot influence any city beyond 1000 kilometers.

3.3 Structural estimation

3.3.1 Methodology

How fast ideas spread in the affiliation network depends crucially on the link activation probability α. Estimating this probability is difficult because we do not observe the spread of ideas directly; we only observe some outcomes of the ideas, after some time. These outcomes are related to exposure to ideas through statistical models, described below. We will use two of these outcomes to construct a confidence interval for α based on the profile of their likelihood. This setup resembles indirect inference or simulated maximum likelihood (Smith, 2008), where one chooses α to maximize the likelihood of the observed data, given the structural model's output. In practical terms, we take the following steps.

1. We fix α at gridded values over $[0, 1]$ at intervals of 0.05.
2. For each value of α, we simulate the spread of two key ideas through the affiliation network and the epidemiological model.
3. We compute exposure of university cities S_t^k, given by equation 3.3, to these ideas.
4. We estimate two auxiliary statistical models correlating the simulated exposure with observed outcomes (detailed below).
5. We record the log-likelihoods as the sum of the individual likelihoods at each point:
$$\ell_{\text{total}}(\alpha) = \ell_1(\alpha) + \ell_2(\alpha)$$
This approach treats the two outcomes as conditionally independent given α.
6. We maximize the combined log-likelihood
$$\hat{\alpha} = \arg\max_{\alpha} \ \ell_{\text{total}}(\alpha).$$
7. We construct a likelihood-based confidence interval for α as the set of values for which the log-likelihood is not "too much worse" than the maximum.
$$2\left[\ell_{\text{total}}(\hat{\alpha}) - \ell_{\text{total}}(\alpha)\right] \leq \chi^2_{1,0.95} \approx 3.84$$
This gives a 95% confidence interval for the link activation probability α based on both outcomes. We let this interval be: $[\alpha_{\text{low}}, \alpha_{\text{high}}]$
8. We take the lower bound for the subsequent analysis: $\alpha = \alpha_{\text{low}}$. It gives the most "conservative" link activation probability that is still consistent with the combined data under the likelihood ratio criterion. In taking the lower bound, we are conservative, in the sense that we minimize the risk of overemphasizing the role of face-to-face interactions within institutions in the spread of ideas.

We now describe the two auxiliary models referred to above in item 4.

3.3.2 Auxiliary model 1: Botanical Realism and botanic gardens

During the Scientific Revolution there were major advancements in botany, and it grew from being primarily a descriptive field into a more systematic and experimental science—a shift we call "Botanical Realism". A key figure in this transition was Leonhart Fuchs, a German physician and botanist. He is best known for his book *De historia stirpium commentarii insignes*, which translates to "Notable commentaries on the history of plants." Printed in Basel in 1542, this work laid the foundation for modern botany. Fuchs not only provided visual representations of 511 plant species: he also included his own critical observations on their uses and characteristics, highlighting differences from ancient texts (Applebaum, 2003).[15] Fuchs was based in Tübingen for the majority of his life, were he taught medicine and botany at the local university between 1535 and 1566 (Conrad, 1960). Prior to that he was a professor at the University of Ingolstadt from 1522 to 1533 (Schwinges & Hesse, 2019). Despite his fame, he was not a mobile scholar and he declined prestigious teaching offers from Denmark and Italy (Applebaum, 2003).

In this first empirical assessment, we examine the potential correlation between exposure to Botanical Realism and the establishment of botanic gardens. We calculate exposure to original botanical ideas in the following way: the diffusion of the idea begins with Leonhart Fuchs and spreads to his colleagues at the University of Tübingen, and then extends further through mobile scholars—those who were affiliated with multiple institutions throughout their lifetimes. Using our epidemiological approach, we average the D simulation outcomes to model how these ideas spread across European institutions between 1500 and 1800 (remember each simulation will differ, because the probability of transmission is less than one).

We use the sample of cities that hosted a university between 1600 and 1800, as recorded in our database (De la Croix, 2021), resulting in a total of 185 university cities.[16] For this first experiment, we also gathered information on the existence and founding dates of European botanic gardens from Montreal Botanic Garden (1886). Figure 3.16 (in Appendix 3.C.2) illustrates this sample of cities along with their exposure to Botanical Realism in 1600, 1700, and 1800.

In what follows, we analyze the probability that each university city would host the creation of a botanic garden. Specifically, we are interested in analyzing whether this probability is affected by the institutions of the city being exposed to Botanical Realism. We estimate a Cox proportional hazard model with different

[15] More contextual information is available in Appendix 3.B.1.

[16] Of these, 182 cities had universities that remained operational after 1600, while the cities of Budapest, Palencia, and Bratislava hosted universities prior to 1600, but these universities did not survive beyond that period.

levels of α. Following the Cox model, the probability $h(t)$ of building a botanic garden in a city with fixed characteristics x and time varying characteristics $y(t)$ and $z(t)$ changes with the survival time t according to

$$h(t) = h_0(t)\exp(x\beta + y(t)\gamma + z(t)\zeta) \tag{3.6}$$

where $h_0(t)$ is the baseline hazard. For the time invariant regressors x, we use initial population in 1500 and the Euclidean distance from Tübingen. In a simple gravity model of diffusion, the distance from Tübingen captures the general effect of the invention, and its spread through pathways other than our affiliation network. The regressor of interest is the time-varying exposure $y(t)$. This exposure counts the number of botanists, physicians, and scientists[17] exposed to Botanical Realism at time t. The exposure used here refers to S_t^k in equation 3.3, and it is varying every year. Institutional exposure also takes into account the publication output of an institution's scholars. For the time-varying control $z(t)$, we introduce "non exposure" to Botanical Realism to capture scientists who were susceptible to exposure but were not actually exposed. This coefficient, ζ, allows us to control for alternative pathways—such as different scientific ideas or the general orientation of the faculty—through which botanic gardens might have been established. This control, formally called $\check{S}_t^k$, is computed as follows:

$$\check{S}_t^k = \sum_s \underbrace{q_s}_{\text{quality}} \left(\underbrace{I(s \in V(k,t))}_{\text{membership}} \; \underbrace{I([\bar{i}_s]_{t'} = 0)}_{\text{exposure}} \right) \tag{3.7}$$

where the individual exposure $[\bar{i}_s]_{t'}$ is set to zero to capture only those scholars who were not exposed, but who could have been, given their field of study and the timing of their presence at the institution. By proceeding in this way, we obtain a dynamic measure of "exposure" for non-exposed scientists that mirrors S_t^k.

We consider the time period 1500-1793. The first botanic garden is observed in 1520 in Pavia, Italy; Fuchs' invention takes place in 1542 and the affiliation data run through 1793. In some cities, no botanic garden was built during this period, and we assume that the garden was created after the censoring data cutoff of 1793, according to the construction of the Cox model.[18] To compute the "risk" of getting a botanic garden, we simulate the cumulative hazard function using the estimated vector of parameters $\hat{\beta}$. In turn, the probability of experiencing the creation of a garden is just the inverse of the computed survival. We use the

[17]These scholars work in fields such as medicine, botany, mathematics, physics, chemistry, astronomy, and applied sciences such as agronomy and engineering.

[18]The fact that botanic gardens were established in only 59 out of 185 cities before 1793 limits the implementation of our Cox model. Specifically, we cannot stratify the fitting routine by city, as too many cities have no event. To address this issue, we cluster the standard errors at the city level.

Nelson-Aalen estimator to compute the baseline hazard: this estimator sums the hazards over the cities still at risk. Using t_i to indicate the different years in which a garden was created, we obtain the expected number of events as follows:

$$E(g_i) = \sum_{j:t_j>t_i} \hat{h}_0(t_i) \exp(x\hat{\beta} + y(t)\hat{\gamma} + z(t)\hat{\zeta}) \tag{3.8}$$

where g_i is the number of events at a specific time t_i and the sum only considers cities still at risk at that specific time t_i (i.e., cities without a botanic garden at t_i). We replace the expected number of events $E(g_i)$ with the actual number of gardens created and we obtain the estimate of the baseline hazard $\hat{h}_0(t_i)$.

By construction, the Cox model assumes time to be continuous, meaning that each botanic garden should have been created one at a time, with no years in which gardens were simultaneously created in more than one city. However, in our sample there are six instances of two botanic gardens being created in the same year–occurrences known in the literature as “ties” or “tied events”. We use the Efron method to manage this in the likelihood calculation and to better clarify the order of events. This method assumes the tied events occurred in small groups and evenly distributes the risk across cities within the same group. While this is an approximation, it efficiently computes the partial likelihood. To ensure robustness, we also apply the exact method, which computes partial likelihoods by systematically evaluating all combinations of tied events. The exact method is the most accurate but it is less flexible (since it does not allow for validation tests) and more computationally demanding. Still, it produces results very similar to the Efron method.

Table 3.2 shows the results for the Cox proportional hazard models at different level of α. We run similar specifications for every α between 0 and 1 at an interval of 0.05 to obtain the log likelihoods which are the first necessary component for our structural estimation.

3.3.3 Auxiliary model 2: Mathematical Astronomy and astronomical observatories

In the 15th and 16th centuries, growing interest in experimental science led astronomers to challenge Ptolemaic models and refine them through observation and mathematics. This shift marked the start of the astronomical revolution, with advances in trigonometry, geometry, and the use of decimals, and a new focus on underlying physical causes rather than mere description.

A key figure in this astronomical revolution was Regiomontanus (which was a pseudonym of Johannes Müller). His mastery of Greek and mathematics

TABLE 3.2: Cox Proportional Hazards Model – Botanical Realism and botanic gardens.

	Dependent variable: Hazard rate of botanic garden founding					
Alphas	0	0.1	0.3	0.5	0.7	1
(ihs) Exposure to		0.079	0.225**	0.223***	0.222***	0.224***
Bot. Real. S_t^k		(0.287)	(0.090)	(0.073)	(0.069)	(0.068)
(ihs) Non exposure to	0.397***	0.505***	0.529***	0.545***	0.551***	0.554***
Bot. Real. $\check{S}_t^k$	(0.070)	(0.074)	(0.077)	(0.078)	(0.079)	(0.079)
(ihs) Distance to	-0.138**	-0.237***	-0.205***	-0.192***	-0.187***	-0.184***
Tübingen	(0.054)	(0.060)	(0.058)	(0.057)	(0.057)	(0.057)
Log Likelihood	-295.234	-295.529	-293.516	-292.034	-291.396	-291.037
(ihs) Pop in 1500	YES	YES	YES	YES	YES	YES
Observations	54390	54390	54390	54390	54390	54390

Note: *p<0.1; **p< 0.05; ***p<0.01. Robust standard errors are reported in parentheses.
Similar models run for every α between 0 and 1 at an interval of 0.05. Here, we only report results for 0, 0.1, 0.3, 0.5, 0.7, and 1.
"Bot. Real." refers to "Botanical Realism". All the variables are transformed in inverse hyperbolic sine (ihs). Distance to Tübingen is computed as Euclidean distance. All models include (ihs) population in 1500.

enabled him to study the original works of Ptolemy and other ancient thinkers. At the University of Vienna, around 1454, he and his mentor, Georg Peurbach (1423 – 1461) began collaborating on new methods for solving plane and spherical trigonometry problems, including the use of sine and tangent functions. Regiomontanus also created extensive trigonometric tables with values calculated to decimal units, which remained influential for centuries. As such, he can be considered a pioneer of Mathematical Astronomy (Applebaum, 2003).[19] Regiomontanus published *Theoricæ novæ planetarum*, his collaboration with Peurbach, in 1472 after Peurbach's death.

In this second empirical assessment, we examine the correlation between exposure to Mathematical Astronomy and the creation of astronomical observatories. We posit that advances in trigonometric methods create demand for better places, buildings, and instruments that can manage more precise astronomical observations. With these observations, scientists can determine with more accuracy the dates of equinoxes, solstices, and other celestial events. Regiomontanus himself opened an instrument shop that specialized in building and printing works related to Mathematical Astronomy (Applebaum, 2003). We collected the names and foundation dates of observatories from Howse (1986).

[19] More details are available in Appendix 3.B.2.

The computation of exposure to Mathematical Astronomy follows the same methodology that we used for Botanical Realism. The idea starts with Regiomontanus, who shared it with his colleagues in Vienna, Bratislava, Padua, and Rome, and reaches other institutions through mobile scholars. After averaging the simulation outcomes, we calculate the S_t^k yearly exposure of each institution to Mathematical Astronomy over time between 1500 and 1793. Figure 3.18 (in Appendix 3.C.3) depicts this sample of cities along with their exposure to Mathematical Astronomy in 1600, 1700, and 1800. Only scientists are considered to have been exposed–scholars working in fields such as mathematics, logic, physics, chemistry, biology, astronomy, geography, and botany. As before, we account for the quality of their publications. Finally, we obtain the institutional exposure to Mathematical Astronomy.

We estimate the probability of each university city obtaining an observatory using a Cox Model similar to that used for Botanical Realism. The technical details, including equations 3.6 and 3.8, remain the same. The only difference lies in the covariates included: we replace distance from Tübingen with distance from Vienna. We focus on the same period, 1500-1793, since the first observatory was established in Kassel in 1560, and we date the spread of Regiomontanus' ideas as starting in 1454 and persisting up to 1793, the last date in our timeframe. 52 cities in the sample saw the creation of an observatory before the censoring date of 1793, while the remaining cities never saw one. However, the Cox Model assumes by construction that these cities will eventually have an observatory after 1793.[20] Again, we encounter tied events in the establishment of observatories. Specifically, there are ten years in which two observatories were constructed simultaneously, and one year (1790) in which three observatories were constructed. We proceed as in Section 3.3.2: the main results are computed with the more parsimonious Efron method, and we confirm the results with exact method.

Table 3.3 shows the results for the Cox proportional hazard models at different level of α. We run similar specifications for every α between 0 and 1 at an interval of 0.05 to obtain the log likelihoods which are the other necessary part for our structural estimation, together with the log likelihood from Table 3.2.

3.3.4 Results

The joint likelihood $\ell_1(\alpha) + \ell_2(\alpha)$ attains a maximum at -539.42 with $\hat{\alpha} = 1$. The lowest admissible likelihood (i.e. not significantly different from its maximum) is

$$\ell_{\text{total}}(\alpha_{\text{low}}) = \ell_{\text{total}}(\hat{\alpha}) - 3.84 = -543.26.$$

[20]As with the botanic garden model in Section 3.3.2, we address this issue clustering the standard errors at the city level; otherwise, the maximization does not converge to a finite likelihood.

TABLE 3.3: Cox Proportional Hazards Model – Mathematical Astronomy with Non Exposure.

	Dependent variable: Hazard rate of observatory founding					
Alphas	0	0.1	0.3	0.5	0.7	1
(ihs) Exposure to		0.314***	0.293***	0.281***	0.281***	0.288***
Math. Astr. S_t^k		(0.068)	(0.057)	(0.056)	(0.054)	(0.053)
(ihs) Non exposure	0.374***	0.546***	0.562***	0.540***	0.543***	0.546***
to Math. Astr. $\check{S}_t^k$	(0.063)	(0.094)	(0.093)	(0.097)	(0.096)	(0.096)
(ihs) Distance	-0.151***	-0.169***	-0.158***	-0.153***	-0.152***	-0.151***
to Vienna	(0.044)	(0.044)	(0.043)	(0.043)	(0.043)	(0.043)
Log Likelihood	-250.707	-250.243	-248.470	-249.257	-248.944	-248.379
(ihs) Pop in 1500	YES	YES	YES	YES	YES	YES
Observations	54390	54390	54390	54390	54390	54390

Note: *p<0.1; **p< 0.05; ***p<0.01. Robust standard errors are reported in parentheses.
Similar models run for every α between 0 and 1 at an interval of 0.05. Here, we only report results for 0, 0.1, 0.3, 0.5, 0.7, and 1.
"Math. Astr." refers to "Mathematical Astronomy". All the variables are transformed in inverse hyperbolic sine (ihs). Distance to Vienna is computed as Euclidean distance. All models include (ihs) population in 1500.

This value is approached for $\alpha \approx 0.25$ with $\ell_{\text{total}}(\alpha_{\text{low}}) = -542.81$. Therefore, $\alpha = 0.25$ is a conservative estimate of the link activation probability.

Before interpreting the results with $\alpha = 0.25$, we first verify that the proportional hazards assumption holds for this level of alpha. This assumption requires that the hazard ratios for the exposure and other covariates remain constant over time. One test for proportionality calculates the scaled Schoenfeld residuals for each covariate and correlates them with time. The assumption is validated if the correlation is not statistically significant (Schoenfeld, 1982). In our preferred specification with alpha = 0.25 (Table 3.2), we compute the correlation between the hazard ratios (i.e., scaled Schoenfeld residuals) and time, both individually for each variable and jointly at the global level. The individual hazard ratio of "(ihs) Exposure to Botanical Realism" shows no correlation with time. Additionally, the global correlation for Column (3) has a p-value of 0.048, indicating that the joint correlation between the hazard ratios and time is significantly different from zero at the 5% level. However, we are confident in validating the proportionality assumption after analyzing the plot of this correlation in Figure 3.20a in the Appendix.

Having tested for the suitability of the Cox Proportional Hazard Model, we can interpret the results. For a sound interpretation, we focus on the hazard ratios, which are computed by exponentiating the coefficients (similar to interpreting odds ratios in logistic regression). The hazard ratio of "Exposure to Botanical Realism" is 1.25 ($exp^{0.221}$), implying that a city with an exposure of 1 to Botanical Realism has a 25% higher probability of having a botanic garden, compared to a city with zero exposure. Remarkably, the coefficient of exposure is significant even when we control for all the other possible pathways of idea diffusion, captured by the distance to Tübingen, and also when we introduce an additional control, "(ihs) Non exposure Botanical Realism $\check{S}_t^k$," which captures the presence of scientists at the institution who were not exposed to the idea of Botanical Realism.

It is challenging to picture the size of these coefficients due to the variation in exposure over time. To address this, in Figure 3.17 (in Appendix 3.B.1) we plot the probability of a city having a botanic garden for different levels of constant exposure over time. This figure represents an "average city", with an average population in 1500 ((ihs) Population in 1500 μ = 2.8), located at an average distance from Tübingen (i.e., (ihs) Distance to Tübingen μ = 7.04), with an average "(ihs) Non exposure" = 0.57. Figure 3.17 shows that if a city maintains full exposure (i.e., a maximum exposure of 5.75), over the entire 294-year period, it will follow the dot-dashed line and get a garden with 100% probability just before 1550. In contrast, for cities with lower exposure levels, such as 0.31 (dashed line, representing the mean of exposure), 1 (dotted line) and 0 (solid line), a 100% probability of hosting a garden is never reached. For cities with an exposure of 1, the probability of having a garden reaches around 73% only towards the end of the period, around 1793. Conversely, cities with an average exposure of 0.31 will have almost a 50% chance of getting a garden, following the dashed line. The solid line, representing no exposure (0), serves as a baseline, showing that an average city will have about a 38% probability of getting a garden by 1793. Overall, these results demonstrate that being exposed to Botanical Realism exponentially increases the probability that an average city will get a garden: from 38% probability, it jumps to 50% at the mean level of exposure and reaches 75% with a constant exposure of 1.

For Mathematical Astronomy and astronomical observatories, we validate the suitability of the Cox Proportional Hazard Model in the same way as we did for Botanical Realism, following Schoenfeld (1982). We find that the correlation between "Exposure to Mathematical Astronomy" and time is not significant when taken individually. On the other hand, the joint correlation of all the covariates with time is only slightly statistically different than zero at the 10% level (global p-value = 0.098). We are confident that the proportionality assumption is validated, especially after analyzing Plot 3.20b of the joint correlation in Appendix 3.C.3.

For the interpretation of the results we again need to consider the hazard ratios. The hazard ratio of "Exposure to Mathematical Astronomy" is 1.35 ($exp^{0.297}$), indicating that a city with an exposure of 1 to Mathematical Astronomy has a 35% higher probability of getting an observatory compared to a city with 0 exposure. Again, as in Subsection 3.3.2, controlling for potential alternative pathways of diffusion of Mathematical Astronomy does not undermine the relevance of our main variables of interest. In addition, our coefficient of interest, “Exposure to Mathematical Astronomy”, remains highly significant when we control for the presence of other scientists at the institution who were not actually exposed to the idea (as in equation 3.7). This new control, “Non exposure to Mathematical Astronomy” may also capture that other ideas—orthogonal to Mathematical Astronomy—could have led independently to the creation of astronomical observatories through different channels. Hence, we also show that controlling for it does not diminish the statistical and economic significance of our main measure of exposure.

Figure 3.19 (in Appendix 3.B.2) plots the probability of getting an observatory for different levels of constant exposure, which allows for a more straightforward baseline and for a better interpretation of the size of the coefficients. We took an “average city” with an average population in 1500 ((ihs) Population in 1500 $\mu = 2.8$), at an average distance from Vienna ((ihs) Distance to Vienna $\mu = 7.38$), and with an average “(ihs) Non exposure to Mathematical Astronomy” of 0.42. We can see in the Figure that a city that is never exposed to Mathematical Astronomy will have a probability of approx. 25% of seeing the creation of an observatory by 1793. This can be considered as the baseline for how an increase in exposure to Mathematical Astronomy impact the likelihood of getting an observatory: already with a constant average exposure of 0.5 (i.e., mean exposure), a city will have an almost 32% chance of getting an observatory, and this probability jumps to more than 50% with a constant exposure of 1. Finally, in the extreme case in which a city always has the maximum level of exposure to Mathematical Astronomy, it will have a 100% probability of having an Observatory after 1560, the year that the first observatory in our sample was created.

3.4 Further empirical assessments

We use the model to extend our empirical analyses beyond the context of the Scientific Revolution and a direct mapping between ideas and outcomes. As with the earlier sections, our analysis is shaped by the availability of outcome data. We examine three distinct cases: a backlash against an idea (the nexus between Scholasticism and Protestantism); an extension of the plague–pogrom nexus that incorporates exposure to anti-Judaic ideas, and the diffusion of a demonstrably false belief that Swedes are descendants of Atlantis.

3.4.1 Scholasticism and Protestantism

Scholastic theology is an approach to theological questions that uses logical analysis and systematic reasoning, influenced by ancient Greek philosophers.[21] It is more a paradigm than a single idea. Petrus Lombardus is often recognized as an early proponent and influential figure in the scholastic tradition. According to Genet (2019) and Herbermann (1913) he taught at what would become the University of Paris from 1145 to his death in 1160. Mazzetti (1847) claims that he was at the University of Bologna in about 1150.

Petrus Lombardus' primary work is the *Sentences*. Completed in the mid-12th century, the *Sentences* cover key theological topics such as the nature of God, creation, the Trinity, grace, and sacraments. The *Sentences* became a foundational text for theological education in medieval universities and was the starting point for many scholastic theologians who followed, including Thomas Aquinas, who wrote extensive commentaries on it.

Martin Luther (1483–1546), the 16th-century German monk and theologian who sparked the Protestant Reformation, was initially trained in the scholastic tradition and engaged with its methods. But as his personal spiritual crisis deepened, he became increasingly critical of many aspects of the Catholic Church's theology—including Scholasticism's emphasis on human reason.

Luther laid out his objections in a striking document, the *Disputatio contra scholasticam theologiam* (1517), a series of 97 theses. In it, he made provocative claims such as: "No syllogistic form is valid when applied to divine terms," and "...the whole Aristotle is to theology as darkness is to light" (theses 47 and 50, respectively).

A major grievance that fueled the rise of Protestantism was the desire to reform theological teachings and Church practices that, in the Reformers' view, were not grounded in Scripture. Scholastic theology—especially in its later form known as nominalism—was a central target. The historian Chaunu (2014) argues that this style of theology, which emphasized logic and abstraction, distanced ordinary believers from their faith and ultimately left them receptive to the message of the Reformation. From this we draw our hypothesis that the adoption of Protestantism was a backlash to (exposure to) Scholasticism. This view is rarely made explicit in the scholarly literature, but it underlies much of the Reformation's intellectual context.[22]

We highlight three key ways in which Luther offered a clear and spiritually compelling response to the crisis of salvation induced by scholastic theology—each emphasizing *direct access to God* by *faith* rather than through a rational

[21] See an example in Appendix 3.C.3.

[22] The theologian Barrett (2023) argues that it was the degeneration of Scholasticism in the later Middle Ages that was a significant catalyst for the Reformation. This view is not universally accepted. For example, Cross (2024) sees substantial continuity between Luther and late Scholasticism.

merit-based system (Chaunu, 2014).

(A) The Catholic Church taught that salvation comes through both faith and works, a position formalized in scholastic doctrines of grace and merit. Luther rejected this, arguing that Scripture teaches that salvation comes by faith alone (*sola fide*), not through human effort or achievement.

(B) Catholic theology placed Scripture and Church tradition—along with papal authority—on equal footing. This framework relied on a scholastic synthesis of Aristotelian philosophy and ecclesiastical tradition. Luther broke with this, insisting that Scripture alone (*sola scriptura*) is the final authority in matters of faith. He viewed the scholastic approach as placing human reason above divine revelation.

(C) The Catholic Church mediated grace through a complex sacramental system, including the sale of indulgences and a strict divide between clergy and laity. Luther opposed this mediation of grace and denied the special status of clergy rooted in scholastic definitions of ordination and apostolic succession. Instead, he emphasized the "priesthood of all believers" (*sola gratia*).

To test for a negative correlation between the exposure to Scholasticism and the rise of Protestantism, we propose the following experiment. We feed the idea of Scholasticism to Petrus Lombardus. The idea would spread to his colleagues in Paris, and then beyond, thanks to the mobility of scholars. We simulate the spread of this idea in Europe, using the epidemiological approach described in Section 3.2.4, all the temporary networks built from our data, and our alpha of 0.25.

Averaging the outcomes of many simulations, we compute the $\tilde{S}_t^k$ exposure of each university to Scholasticism in 1508, the year Luther started teaching at the University of Wittenberg. We measure exposure by counting the number of theologians exposed to Scholasticism in the previous 30 years, by university. Each theologian is weighted by the importance of his publication output. We annualize this exposure by dividing by the 30 years over which we counted the active theologians. We then obtain the exposure to Scholasticism of each university in the network, using equation 3.4.[23] Figure 3.6 depicts this level of exposure. We can further compute the exposure to this paradigm for European cities even if they do not host any university, as in equation 3.5. We do this by computing the distance between each city in our sample and each university in our network, and summing up universities' exposure at the city level weighted by the inverse of distance, w_{ck}.

The sample of cities used to compute exposure is taken from Rubin (2014), which provides a database of over 800 European cities and classifies them as

[23] It is important to note the difference between this "static" exposure and the "yearly" exposures used in the previous two empirical assessments.

Catholic or Protestant based on the dominant religion in three different years: 1530, 1560, and 1600. For this paper, we use the same classification, detailed in Appendix A of Rubin (2014).[24]

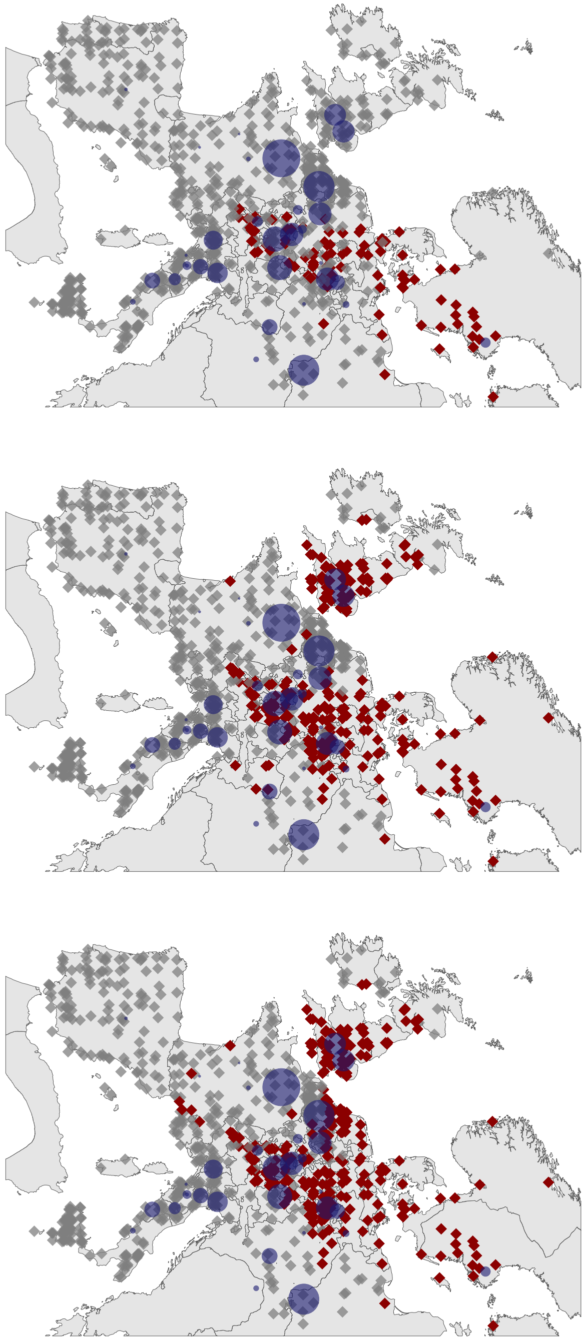
Figure 3.6 illustrates the sample of cities and the spread of Protestantism. Cities that remained Catholic are shown in grey, while those that became Protestant are marked in red. The blue bubbles represent universities' exposure in 1508, as described earlier. The figure already suggests a possible positive correlation between exposure to Scholasticism, and the likelihood that a city would reject Scholasticism in favor of Protestantism. However, it is important to note that in Italy, institutional exposure to Scholasticism was weaker than in Northern Europe. This is because in Italy scholastic ideas were primarily discussed in monasteries and convents, rather than within universities. In Spain, the scholastic paradigm developed later, mainly in the 16th century, emerging from the works of Francisco de Vitoria (c. 1483 - 1546).

We employ a linear probability model to better estimate the correlation between exposure to Scholasticism and the likelihood that a city became Protestant in 1530, 1560, and 1600. For this experiment, we cannot use a Cox proportional hazard model because in some European regions, such as England and Scotland, the shift towards Protestantism was a top-down decision such that all the cities became Protestant on the same day. This would create too many ties, violating the Cox model's assumption of time being continuous.

Columns (1)-(3) in Table 3.4 show the results. The key variable of interest is *Exposure to Scholasticism* S^c_{1508}, whose estimated coefficient is consistently positive and statistically significant in 1560 and 1600. Columns (1)-(3) show the variable of interest without fixed effects. Here, we only control for the presence of universities in 1500 to show that *Exposure to Scholasticism* S^c_{1508} is not directly substituted by the simple presence of a university. This means that it is not sufficient that a city has a university: it must also be exposed to Scholasticism either directly via the university in that city, or indirectly, by being close to other exposed cities. In Appendix 3.C.3 Table 3.12, we include additional controls and fixed effects.

Using the estimated coefficients in Table 3.4, Columns (1)-(3), we can assess the magnitude of the relationship between a city's exposure to Scholasticism and its likelihood of adopting Protestantism. Given Heidelberg's position as the place most exposed to Scholasticism, we can estimate how the probability of becoming Protestant might have changed for other cities in our sample, had they experienced a similar degree of exposure. For example, considering Barcelona in Spain, which is in the lowest quartile of the exposure distribution, we find that its probability of becoming Protestant would have increased by 12.3% in

[24]We updated Rubin's data to better account for the fact that in France and the Low Countries (modern-day Belgium), several cities adopted Protestantism temporarily before being reconquered by Catholic forces. See Appendix 3.C.3.



(A) Protestant cities in 1530

(B) Protestant cities in 1560

(C) Protestant cities in 1600

FIGURE 3.6: Blue bubbles represent the exposure to Scholasticism 30 years prior to 1508, $\alpha = 0.25$ and $D = 5{,}000$. Protestant cities are the red diamonds, and Catholic cities are the grey diamonds. Data on cities' religion are taken from Rubin (2014) and updated as in Appendix 3.C.3.

1530, by approximately 37% in 1560 and by 49% in 1600. Looking at a city in the second quartile, such as Copenhagen in Denmark, the increase in probability would be 9.6% in 1530, 28.7% in 1560, and 38.3% in 1600. For Bologna in Italy, which falls in the middle of the third quartile, the increase would have been about 5.8% in 1530, approximately 17% in 1560, and 23% in 1600. In contrast, a city in the same quartile as Heidelberg, such as Leuven in Belgium, which has a difference in exposure of only 12.4 absolute points compared to Heidelberg, would have a much lower increase in probability: 1.2% in 1530, approximately 4% in 1560, and 5% in 1600.

TABLE 3.4: Linear Probability Model - Exposure to Scholasticism in 1508 and cities' probability of becoming Protestant in 1530, 1560, and 1600.

	Protestant in			Protestant in		
	1530 (1)	1560 (2)	1600 (3)	1530 (4)	1560 (5)	1600 (6)
Exposure to Scholasticism S^c_{1508}	0.001 (0.001)	0.003*** (0.001)	0.004*** (0.001)	0.0005 (0.001)	0.005*** (0.002)	0.006*** (0.002)
Presence of university in 1500	-0.034 (0.027)	-0.075 (0.051)	-0.130** (0.054)	-0.044 (0.027)	-0.018 (0.045)	-0.056 (0.047)
Non exposure to Scholasticism $\check{S}^k_{1508}$				0.006 (0.005)	-0.034 (0.024)	-0.043** (0.020)
Observations	867	867	867	867	867	867
Adjusted R^2	0.016	0.072	0.127	0.018	0.116	0.194
Log Likelihood	-201.02	-500.48	-515.10	-199.68	-478.98	-480.13

Note: *p<0.1; **p< 0.05; ***p<0.01. Robust standard errors clustered by territory are reported in parentheses. A constant term is included in all regressions. Dependent variable "Protestant" takes value 1 if the city is Protestant in 1530, 1560, 1600, respectively. Data on cities' religion taken from Rubin (2014) and updated as in Appendix 3.C.3.
"Presence of university in 1500" is a dummy variable taking value 1 if our database shows the city having a university in 1500 (De la Croix, 2021). "Exposure to Scholasticism S^c_{1508}" and "Non Exposure to Scholasticism $\check{S}^k_{1508}$" are computed as in equations 3.4 and 3.7, respectively.

In Columns (4)-(6) we show how the correlation between the *Exposure to Scholasticism* S^c_{1508} and the probability of becoming Protestant does not change much when we control for the "Non exposure to Scholasticism $\check{S}^k_{1508}$", capturing the theologians susceptible to being exposed to Scholasticism but who are not, as in the case of Botanical Realism and Mathematical Astronomy. It is remarkable that also when introducing this additional control, our coefficients of interest in Columns (5)-(6) remain highly significant and with a slightly larger magnitude than Columns (2)-(3). We interpret this as an indication of robustness: controlling for both the simple presence of a university and the "Non exposure", we still see a positive and significant correlation between the latter and the probability that the city would become Protestant. This reinforces our initial hypothesis of a backlash to exposure to Scholasticism.

3.4.2 Anti-Judaism and the Persecution of Jews

Our model can propagate not only good ideas, but also bad or even false ones. One such idea is anti-Judaism. The availability of data on Jewish persecutions from Anderson, Johnson, and Koyama (2017) and Jedwab, Johnson, and Koyama (2019) allows us to apply the same approach as in Section 3.4.1 to analyze the spread of anti-Judaism.

The roots of anti-Judaism run deep and are widely discussed in the literature. Scholasticism, once again, played a role in rationalizing prejudice against Jews. Thomas Aquinas (1225–1274)—one of the most influential scholars at the University of Paris and the intellectual heir of Lombardus—endorsed many prevailing medieval Christian views about Jews. He supported the idea that Jews should live in subjugation as a reminder of their supposed rejection of Christ. In the *Summa Theologiae*, he also discusses Jews in ways that reinforce their marginalization.

The next generation of scholastic theologians, such as John Duns Scotus (c. 1266–1308) and William of Ockham (c. 1287–1347), contributed to the broader scholastic discourse that pathologized Judaism as a theological error (see Abulafia (2011), which provides a substantial discussion of how scholastic theology and legal reasoning shaped Christian–Jewish relations between 1000 and 1300). While these theological ideas do not explicitly advocate persecution, they may have interacted with the mechanisms identified by Anderson, Johnson, and Koyama (2017) and Jedwab, Johnson, and Koyama (2019). In Anderson, Johnson, and Koyama (2017), cold temperatures are shown to increase the probability of persecution of Jewish communities. Jedwab, Johnson, and Koyama (2019) extend this insight, showing that negative shocks more broadly—particularly plagues—raise the likelihood of minority persecution.

To study the correlation between yearly *Exposure to Scholasticism* S_t^c and the probability of violent acts against Jews, we use data from Anderson, Johnson, and Koyama (2017) and Jedwab, Johnson, and Koyama (2019).[25] Building on this literature, we hypothesize that the effect of negative shocks on the likelihood of Jewish persecutions is amplified in cities with greater *Exposure to Scholasticism* S_t^c. The underlying intuition is that when local priests—shaped by scholastic teachings—disseminate anti-Judaic arguments, they create conditions

[25]We thank the authors for sharing the most recent (and still unpublished) version of the pogroms and persecutions data.

that make communities more likely to scapegoat Jews in times of crisis. To test this hypothesis, we augment the empirical framework of Anderson, Johnson, and Koyama (2017) by including an interaction term between our measure of scholastic exposure and the incidence of plague outbreaks.

TABLE 3.5: Linear Probability Model - Yearly Exposure to Scholasticism and cities' probability of persecution of Jews.

	Persecutions	
	Replication (1)	S_{ct} x Plague (2)
$Temperature_{c,t-1}$	-0.467***	-0.496***
	(0.125)	(0.129)
Plague	5.100**	-0.719
	(2.149)	(1.274)
Exposure to Scholasticism S_{ct}		0.025
		(0.052)
Exp. to Scholasticism S_{ct} x Plague		3.621***
		(1.131)
Controls	YES	YES
City Fixed Effects	YES	YES
Observations	273,879	273,879
R^2	0.013	0.015

Note: *p<0.1; **p< 0.05; ***p<0.01. Standard errors clustered at the climate grid level in parentheses. City fixed effects are always included.
Column (1) replicates Anderson, Johnson, and Koyama (2017) specification (2) Table 3, p.940. Coefficients are multiplied by 100 to represent percentage points.
Controls include a slope variable for the 10 years surrounding the Black Death and a measure of population density as in Anderson, Johnson, and Koyama (2017).

Table 3.5 presents the results of this linear probability model. The dependent variable, $Persecutions$, takes the value 1 when either an expulsion or another violent act against Jews occurred in a given year, following the definition in Jedwab, Johnson, and Koyama (2019). As in both Anderson, Johnson, and Koyama (2017) and Jedwab, Johnson, and Koyama (2019), we restrict the analysis to cities with a documented Jewish presence.

Column (1) replicates specification (2) of Table 3 in Anderson, Johnson, and Koyama (2017, p. 940). The estimated coefficient on lagged temperature is nearly identical: we find that a one-degree decrease in temperature increases the probability of Jewish persecutions in the following year by 0.467 percentage points, compared to 0.464 percentage points in Anderson, Johnson, and Koyama

(2017). Despite this similarity, our sample differs slightly for two main reasons: (i) our dependent variable is drawn from a more recent version of the dataset used by Anderson, Johnson, and Koyama (2017), and (ii) we eliminate duplicate city entries prior to estimation.

Column (2) introduces an interaction between our yearly *Exposure to Scholasticism* S_t^c at the city level and a dummy variable indicating the presence of a plague, following Anderson, Johnson, and Koyama (2017). Interestingly, neither *Exposure to Scholasticism* S_t^c nor the plague dummy is statistically significant on its own. However, their interaction is both statistically and substantively significant. This suggests that when a theoretical framework exists that portrays Jews as a threat, and a plague occurs simultaneously, the probability of violence against Jews rises significantly. The coefficient on the interaction term is also sizable: during a plague, a one-unit increase in exposure to Scholasticism is associated with a 3.6 percentage point increase in the likelihood of Jewish persecutions.

3.4.3 Finding Atlantis: A True Story of Genius and Madness

We now turn to Olaus Rudbeck's (1630 – 1702) claim that Sweden was the cradle of civilization and the site of the lost city of Atlantis. It is an interesting case to simulate using our model for two reasons. First, it illustrates that ideas–whether accurate or not–can still spread through affiliation networks, which further underscores the crucial role of institutions in shaping intellectual diffusion. Second, it highlights that individuals do not necessarily need to agree with an idea in order to help propagate it.

Rudbeck, who was professor of medicine at the University of Uppsala from 1658 to 1692 (Von Bahr, 1945), claimed that Sweden was in fact the mythical island of Atlantis, and thus the cradle of all ancient civilization. He supported this sweeping theory by drawing connections between the Norse mythology, the Bible, and classical sources. Rudbeck explained his thesis in the book *Atlantica* (also known as Atland eller Manheim), first published in four volumes between 1679 and 1702 (King, 2005). It was written in Latin (vol. 1) and Swedish (vols. 2–4). *Atlantica* was not fully translated into major European languages during the 18th century. The length, complexity, and eccentricity of Rudbeck's arguments likely discouraged publishers. His ideas were seen by many contemporaries as extravagant, although some Nordic nationalist thinkers admired them. Despite its eccentricity, *Atlantica* was referenced and critiqued by various 18th-century thinkers, including two academic scholars, Denis Diderot and Ludvig Holberg. They may have been exposed to Rudbeck's ideas indirectly through the affiliation network, given that they are unlikely to have read his books in Swedish themselves.

Denis Diderot (1713–1784) was a member of the Prussian Academy of Sciences from 1751 (Amburger, 1950). Diderot's radical thinking, conflict with French authorities, and nonconformist personality kept him outside the fold of French academies. In the article *Étymologie* of his Encyclopédie (published over the period 1751–1765), Diderot used Rudbeck's work as a cautionary example of how speculative etymology can lead to erroneous conclusions, critiquing the methodology employed in *Atlantica*. Our simulation reveals that Diderot had a 100% chance of being exposed to Rudbeck's ideas as early as 1751–the year of his election–because he hypothetically encountered numerous members who were already exposed to the idea.

Ludvig Holberg (1684–1754) was a prominent Danish-Norwegian writer and philosopher, and professor at the University of Copenhagen from 1717 to 1754 (Slottved, 1978). He satirized Rudbeck's theories, mocking the idea of Sweden as Atlantis and highlighting the speculative nature of Rudbeck's claims. According to our simulation, the chance Holberg was exposed to Rudbeck's idea is 0 until 1753, when it goes to 22.9%. It gains an additional 49.4% in the following year, before he died. This highlights two features of our approach: first, our exposure is a lower bound, and in this case Holberg might have been acquainted with Rudbeck's work through other means. In other words, even in cases where we know from historical evidence that a scholar engaged with an idea, our model still captures eventual exposure through the network alone. Second, Holberg was never affiliated with an academy, but only with a university, which in our model means it must have taken more time for ideas to reach him. As we will later show, this institutional feature plays a key role in shaping how easily ideas circulate through the network.

3.5 Counterfactual experiments

In this section, we identify the features of the academic network that are more conducive to spreading ideas. We perform two kinds of experiment. We assign ownership of an idea to fictitious inventors, in order to track whether the idea would spread differently in alternative realities. In the second experiment we remove some parts of the network to assess their importance in spreading ideas. We first exclude academies, which were more innovative and more connected institutions than traditional universities, from the network. We also remove institutions from certain geographical areas (the British Isles, the Italian Peninsula, Iberia, and France) to assess the historical importance of each region in fostering scientific progress. This removal alters the network by eliminating edges representing affiliations to these institutions, thereby disconnecting scholars solely affiliated with them. Finally, we remove the Jesuits from the network, who were an important component with their c. 6000 scholars and c. 50 higher education institutions.

3.5.1 Placebo inventors of Botanical Realism

To better understand how the structure of a network influences the speed at which ideas spread, we run counterfactual experiments using Fuchs' Botanical Realism as a case study. In these experiments, we imagine that it was not Fuchs who introduced the new paradigm of Botanical Realism, but another contemporary scientist from a different region of Europe. We simulate the diffusion of this paradigm, still originating in 1542, but emerging in various alternative locations: in Salamanca with Juan Aguilera (1507-1560), in Zaragoza with Gaspard Lax de Sarenina (1487-1560), in Oxford with John Warner (c. 1500-1565), in Louvain with Jeremius Dryvere (1504-1554), in Wittenberg with Andreas Goldschmidt (1513-1559), in Cracow with Mikołaj Mleczko Wieliczki (1490-1559), in Rostock with Jacob Bording (1511-1560), in Montpellier with Antoine Saporta (1507-1573), in Padua with Girolamo Donzellini (1513-1587), from the Royal College of France with Oronce Fine (1494-1555), in Pisa with Realdo Colombo (1510-1559), and in Leipzig with Georg Joachim Porris (1514-1574). Appendix 3.C.3 presents a short biography of each of these scholars. These counterfactual simulations allow us to explore how regional networks and academic hubs would have shaped the spread and influence of Botanical Realism across Europe.

Many of these scholars came from strong intellectual backgrounds and were part of the broader Renaissance shift toward empiricism and direct observation. Several, particularly those with medical training (e.g., Saporta, Colombo, Bording), had practical reasons to study plants carefully and could have contributed to a scientific approach to botany.[26]

Figure 3.7 illustrates the percentage of individuals in medicine and sciences who were exposed to the idea across the twelve simulated scenarios. This simulation offers valuable insights into the diffusion process. Notably, in two cases, the idea fails to spread. Wieliczki in Cracow (professor from 1513 to 1552) and Lax in Zaragoza (professor from 1521 to his death in 1560) lacked other mobile peers for meaningful intellectual exchange, and in these two scenarios the idea does not take hold. In ten other cases, we observe that by the end of the period, nearly all relevant scholars had encountered the idea. It is important to note that these results are averaged over 5,000 simulations, meaning a consistently high rate of diffusion across all simulations. This reflects

[26] That said, Botanical Realism required not just empirical skills but also an interest in plants themselves, which some of these figures might not have had. Fuchs' success came from a combination of his medical background, interest in plants, access to talented illustrators, and the specific intellectual environment in Germany at the time.

how effectively the European intellectual network functioned to disseminate ideas. Regardless of their point of origin, ideas eventually spread throughout Europe in the long run. For instance, Warner's idea remained localized in Oxford for some time, but reached Cambridge and Gresham College by 1650, before spreading further. Similarly, Saporta's concept, originating in Montpellier, reached Basel, Lincei, and Toulouse by 1600, and continued to propagate across Europe afterward. After two centuries, both Bording's and Saporta's ideas had spread at a similar rate to various locations. This demonstrates that, despite differences in individual pathways and speed of diffusion, the overall outcome was the same: the widespread diffusion of ideas across Europe.

This confirms that the diffusion process generated by our model is non-ergodic: the success of an idea remains dependent on its initial conditions – specifically, the network position of its inventor.[27] This implies that the distribution of successful ideas does not converge to a single stationary form, i.e. even as time progresses, the expected success of an idea remains contingent on the initial conditions. It also implies that outcomes across different realizations of the process do not average out over time (in an ergodic process, averaging over time should yield the same result as averaging over different realizations of the process: see Peters (2019) for a history of the idea of ergodicity).

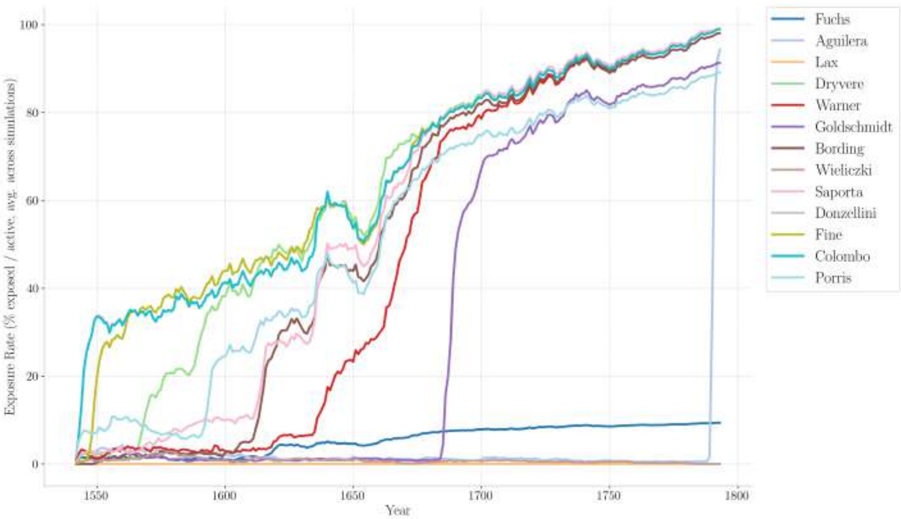


FIGURE 3.7: Exposure rates of active scholars in the network from 1084 to 1793, considering different hypothesized proponents of Botanical Realism.

[27]Appendix 3.I.1 provides a more detailed discussion of this point.

The case of Fuchs (4) is particularly interesting, as the diffusion of his herbal plateaus at 10%, unlike the other simulations where ideas either reach full exposure or fade out entirely. This reflects the fact that, in roughly 90% of the simulations, Fuchs' idea dies out quickly, keeping exposure at zero. In the other 10%, Fuchs' idea spreads successfully, reaching high levels of adoption after a century. This outcome suggests that the spread of Fuchs' herbal hinges on a fragile initial phase, where its survival occurs with a probability of about one tenth.

To understand the European academic network, we can examine the transmission of ideas through the case of Fuchs' Botanical Realism. Figure 3.8 highlights the key individuals and institutions involved. This idea originated at the University of Tübingen, where it thrived for over a century due to the steady presence of scholars in science and medicine. However, the Thirty Years' War effectively closed the university and disrupted this continuity, especially between 1628 and 1634, during its occupation by Imperial (Catholic) forces of the Holy Roman Empire.

Despite this disruption, the idea spread to other institutions via Tübingen scholars who secured positions elsewhere. Jakob Degen taught briefly in Strasbourg (Berger-Levrault, 1890), while Michael Mästlin held a temporary post in Heidelberg (Drüll, 2002). However, these transfers did not result in sustained knowledge transmission: Strasbourg was too small to establish permanent positions in the sciences, and Heidelberg faced the same wartime challenges as Tübingen. Nevertheless, in Strasbourg, the physician Kasper Maliński may have encountered Fuchs' ideas. His subsequent move to the University of Zamość (Kedzoria, 2021) could have carried the concept further.

At Zamość, the mathematician Adrien Van Roomen (also known as Romanus) might have engaged with the idea. Near the end of his life, Romanus became a member of the Accademia dei Lincei, where he potentially reintroduced the concept. Through the Lincei, an informal academy with prominent members such as Galileo and Kepler, the idea could have spread internationally. Thus Adrien Van Roomen, Jakob Degen, Michael Mästlin, and Kasper Maliński are necessary for the survival of the idea. Van Roomen is only the last of a series of key players according to Zenou's (2016) definition, which was developed in the context of criminal networks: "the key player who is the agent that should be targeted by the planner so that, once removed, she will generate the highest level of reduction in total activity" (p. 1403).

This hypothetical trajectory highlights three features of European academia in the sixteenth and seventeenth centuries. First, the dense network of connections ensured that ideas could survive even amid significant disruptions, such as the Thirty Years' War. Marginal institutions, like the University of Zamość, played a crucial role in this resilience. Second, early informal academies, such as the Lincei, were vital for preserving and disseminating ideas across borders. Third, the European academic network was strongly path-dependent and non-ergodic as described in (David, 1985): it was shaped by more or less random historical events “rather than systematic forces" (p.332).

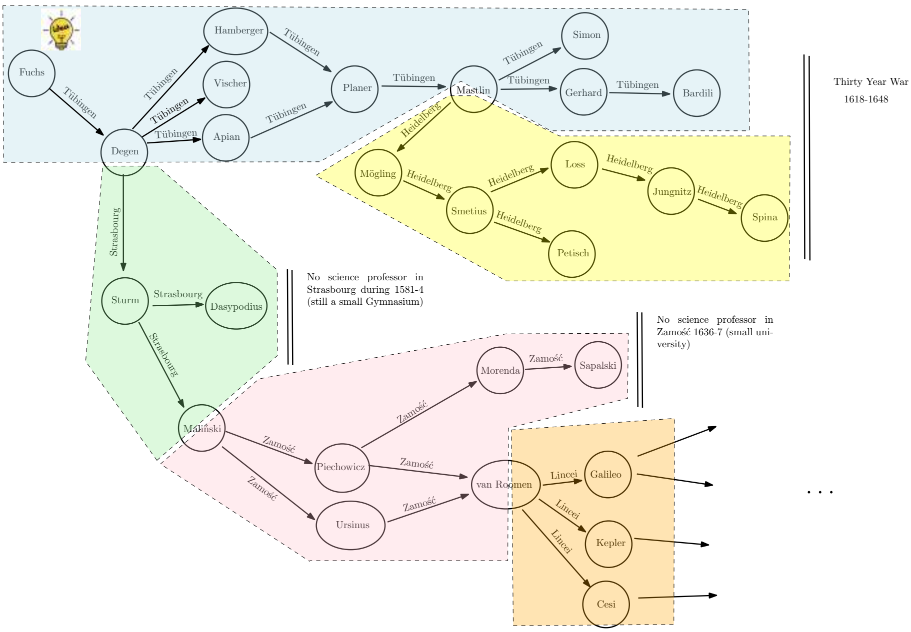


FIGURE 3.8: Botanical Realism path.

3.5.2 Removing components of the network

As explained above, we simulate the spread of ideas in a network stripped of certain components. This gives us better insight into the role of each component. Under this approach, we rewrite equation 3.1 as

$$I_{t+1}^B = B_t I_t^B + I_t^B \tag{3.9}$$

where the new affiliation matrix $B \leq A$ in the Hadamard order (that is, every entry of B_t is less than or equal to the corresponding entry of A_t). Then it is obvious that $I_t^B \leq I_t \ \forall t$, assuming the same initial condition $I_0 = I_0^B$. Indeed, in the new dynamics, there will be fewer exposed persons at every time step, since reducing the number of connections reduces the opportunities for ideas to spread. This follows from the fact that matrix multiplication with a reduced adjacency matrix B_t leads to a weakly lower infection count at every step.

In the stochastic version, as $\mathbb{E}[\Omega^d(B_t)] = \alpha B_t$, we have $\mathbb{E}[\Omega^d(B_t)] < \mathbb{E}[\Omega^d(A_t)]$ in the Hadamard order, and $\mathbb{E}[I_t^B] \leq \mathbb{E}[I_t]$, $\forall t$. This means that, with an infinite number of simulations D, the world with B (fewer contacts) will still have fewer exposures than the world with A, even in the presence of stochastic transmission. But this statement is no longer strictly true in every realization. If we simulate the process many times, the law of large numbers implies that the average outcome of these simulations should converge to the expectation. The required number of simulations D may however be very large, because of strong non linear effects coming from the topology of the network.

First, we let Botanical Realism and Mathematical Astronomy spread over the affiliation network as if academies never emerged, to assess whether academies were key for the diffusion of the ideas of the Scientific Revolution (McClellan, 1985; Pedersen, 1992). To measure diffusion we use equation 3.5, which computes the exposure of any European city to an idea. We use the set of cities in Buringh (2021), excluding those in the Ottoman Empire and in Russia. This leaves us with 1,916 cities. We report quartiles of the distribution of exposure across this set of cities relative to the benchmark.

In reporting the results, we distinguish between two roles of academies. First, they contribute directly to the exposure of nearby cities to ideas – for example, when Greenwich benefits from the presence of the Royal Society in London. Second, they facilitate the diffusion of ideas within the affiliation network by bridging university communities – for instance, when Greenwich benefits from scholars at Oxford and Cambridge, whose intellectual development and connectivity have been enhanced by the Royal Society. In Table 3.6, the line “No direct effect” gives the exposure distribution when the direct effect is shut down. Practically, we keep the vector of the individual exposures I_t^d from the benchmark but we remove academies’ exposure $\tilde{S}_t^k$ from the computation of city exposure S_t^c. The line “No Academies at all” is based on an alternative affiliation matrix where all the edges stemming from academies are removed, and the various measures of exposures are computed with this matrix. The first line for each year “With Academies” represents the baseline exposure distribution, which can be used for comparison.

The following insights can be drawn from Table 3.6: there are very few academies in 1600. The Ricovrati in Padua was only just created (in 1599), and it was mostly literary at that time. The same is true of the Accademia della Crusca (founded in Florence in 1583 to preserve the purity of the Italian language). The Lincei, already mentioned above in the context of its relationship with Romanus and Botanical Realism, was founded in 1603 (Gabrieli, 1989). As a result, academies did not have a big influence on the exposure to Botanical Realism, and both lines remain pretty similar to the benchmark. However, academies already played a role in spreading Mathematical Astronomy, as the exposure distributions drops with respect to the benchmark: the median drops by 23%.

TABLE 3.6: Counterfactual experiment with and without academies.

	Q1	Median	Q3
Botanical Realism			
With Academies in 1600	0	5.18	12.63
No direct effect in 1600	0	5.18	12.63
No Academies at all in 1600	0	5.16	12.58
With Academies in 1650	6.02	16.06	25.42
No direct effect in 1650	3.72	10.04	16.67
No Academies at all in 1650	0	2.51	5.72
With Academies in 1700	12.84	44.42	82.76
No direct effect in 1700	5.57	17.66	28.45
No Academies at all in 1700	0	0	0
With Academies in 1750	29.11	110.34	190.68
No direct effect in 1750	9.13	29.70	53.02
No Academies at all in 1750	0	0	0
With Academies in 1793	78.39	308.81	521.35
No direct effect in 1793	13.02	41.49	75.11
No Academies at all in 1793	0	0	0
Mathematical Astronomy			
With Academies in 1600	0.15	8.67	22.73
No direct effect in 1600	0.14	6.66	16.97
No Academies at all in 1600	0.14	6.45	16.41
With Academies in 1650	20.60	56.15	90.09
No direct effect in 1650	8.38	23.98	39.55
No Academies at all in 1650	0.01	5.38	20.25
With Academies in 1700	45.79	144.73	292.56
No direct effect in 1700	18.21	49.81	77.38
No Academies at all in 1700	0.11	4.57	17.18
With Academies in 1750	124.38	444.15	759.10
No direct effect in 1750	37.99	114.45	200.55
No Academies at all in 1750	11.34	33.27	58.02
With Academies in 1793	358.77	1381.77	2291.75
No direct effect in 1793	51.06	161.96	285.24
No Academies at all in 1793	23.92	60.56	105.60

Summary Statistics of cities' exposure distributions to ideas when [0] academies are fully considered (benchmark) [1] academies have no direct effect on cities but are still present in the network [2] academies have no effect at all.

The indirect effect is smaller, as removing it lead to a larger drop of 25%. Even at this early stage, when academies are few and primarily informal, they contribute to the dissemination of ideas.

By 1650, academies begin to influence Botanical Realism, both as part of the network (due to figures like Romanus) and directly. By this time, both Botanical Realism and Mathematical Astronomy have reached nearly all cities. After 1700 academies are increasingly significant. For Botanical Realism, academies are essential, as indicated by the third row in each scenario dropping to zero. In contrast, while Mathematical Astronomy does not depend strictly on academies for its survival, they play a crucial role in amplifying exposure. By 1793, the absence of academies would lead to a dramatic reduction in exposure to Mathematical Astronomy: in the scenario without academies at all, the median would drop by more than 95%.

We can use this same tool to analyze the role of specific regions or nations. The literature has examined the contributions of each nation to the rise of science and knowledge in Europe. Each country possessed unique characteristics that, when combined, created a fertile environment for intellectual and scientific transformations. For example, Italy laid the foundations with the Renaissance and early scientific methods (Applebaum, 2003); the British Isles drove empiricism and practical applications (Mokyr, 2011a); France spearheaded Enlightenment thinking and institutional science (Ferris, Stella, & Yon, 2010); the Iberic Peninsula advanced economic theory and natural law, and the Holy Roman Empire advanced theoretical frameworks in mathematics and astronomy.

We apply our model to study separately the importance of each nation in spreading and keeping alive each idea. To measure how nations are key for an idea, we construct five counterfactual networks, removing institutions belonging to specific geographical areas: one without the Italian peninsula, one without the British Isles, one without France,[28] one without the Iberian peninsula, and one without the Holy Roman Empire (as defined in Stelter, De la Croix, and Myrskylä (2021). We then simulate the spread of Botanical Realism and Mathematical Astronomy in these five networks, and in the benchmark model.

Results are presented in Table 3.7. Each number is a summary statistic of the exposure S^c_{1793} of the cities' distributions, together with the relative exposure distribution in the benchmark. Without the Italian Peninsula, neither Botanical Realism nor Mathematical Astronomy would have survived. For example, for the latter, the universities of Rome and Bologna were critical hubs, allowing scholars previously exposed to the Mathematical Astronomy in Vienna or Bratislava to continue disseminating it among their colleagues. It appears that the Holy Roman Empire is necessary for Botanical Realism, but not for Mathematical

[28]We exclude from France cities which became French towards the end of the period: Strasbourg (1681), Molsheim (1648), Nancy (1766), Pont-a-Mousson (1766), Nice (1860), Perpignan (1659), Arras (1659), Douai (1667).

Table 3.7: Counterfactual Experiment with and without European regions.

	Q1	Median	Q3
Botanical Realism			
No Italian Peninsula	0	0	0
No British Isles	67.34	240.88	455.91
No France	64.40	280.03	480.19
No Iberic Peninsula	78.39	308.81	521.35
No Holy Roman Empire	0	0	0
Benchmark	78.39	308.81	521.35
Mathematical Astronomy			
No Italian Peninsula	0	0	0
No British Isles	290.39	1021.46	1911.36
No France	234.06	991.08	1668.37
No Iberic Peninsula	340.63	1348.13	2237.27
No Holy Roman Empire	339.30	1282.77	2118.87
Benchmark	358.77	1381.77	2291.75

Summary Statistics of cities' exposure distributions to ideas in five counterfactual networks without a specific European region, and the benchmark in 1793.

Astronomy. No other regions in this simulation were necessary for either idea to spread. For example, both Botanical Realism and Mathematical Astronomy spread throughout Europe even without the British Isles, although the median drops by 22% and by 26%, respectively. Even if the British Isles had a key role in the propagation and implementation of useful knowledge (Hallmann, Hanlon, and Rosenberger (2022) show that British inventors worked on technologies that were more central within the innovation network), it is not case as far as our example of propositional knowledge is concerned.

The same reasoning holds for the other regions: the spread of ideas across Europe is not significantly hindered by the absence of France, and the contribution of the Iberian Peninsula appears particularly limited or even negligible.

Overall, this analysis underscores the resilience of the European network of academies and universities. Even when some parts are removed, the network remains sufficiently dense to sustain the circulation of ideas.

Finally, we focus on the role of Jesuits, and present the simulation results when Jesuits are removed from the network. The Society of Jesus, founded in 1540 by Ignatius of Loyola, was a highly influential religious order in the Catholic Church. Its members underwent rigorous training, including years of spiritual exercise and intellectual formation. To counter Protestantism, Jesuits rapidly established an extensive network of schools, colleges, and universities across Europe (Grendler, 2019) and beyond. In the RETE database, we count 52 Jesuit institutions among the 211 universities and colleges of some renown, and more than 6400 scholars (Jesuit priests)–approximately 8% of all recorded scholars between 1000 and 1800, a figure that rises to 10.9% when the sample is limited to the period after the Jesuit order was founded. Known for their high academic standards, Jesuits taught humanities, sciences, philosophy, and theology. They were also prolific authors: in terms of publications, 5.6% of all VIAF titles associated with RETE scholars have a Jesuit author (6.2% if considering the period after the order was established). Their growing influence led to political tensions and subsequently expulsions from several countries: Portugal (1759), France (1764), Spain (1767), and Naples (1767). In 1773, under pressure from European rulers, Pope Clement XIV suppressed the order, though it survived in Russia and Prussia (and later in the USA, where Georgetown University was founded in Washington DC).

TABLE 3.8: Counterfactual experiment with and without Jesuits.

	Q1	Median	Q3
Botanical Realism			
Benchmark	33.29	126.13	218.01
No direct effect	32.37	121.72	212.79
No Jesuits at all	28.49	107.00	187.43
Mathematical Astronomy			
Benchmark	122.51	437.48	747.74
No direct effect	114.02	417.16	718.26
No Jesuits at all	115.86	423.61	729.94

Summary Statistics of cities' exposure distributions to ideas in the counterfactual network without Jesuits and the benchmark in 1750.

Table 3.8 presents the results for the year 1750 (a few years before the dissolution of the order). When we remove all Jesuit nodes from the network and simulate the spread of Mathematical Astronomy, we observe few differences, except in peripheral regions largely neglected by non-Jesuit institutions (such as Sicily, Andalusia, and Romania). This is somewhat surprising, given the widely recognized contributions of Jesuits to science—particularly astronomy—as evidenced by the many lunar craters named after Jesuit astronomers and all the observatories they built (Udías, 2003). The median drop is between 3% (No Jesuits at all) and 4.6% (No direct effect). One possible explanation for their limited role in the broader diffusion of Mathematical Astronomy is that they operated within a relatively isolated network, separate from the rest of academia.

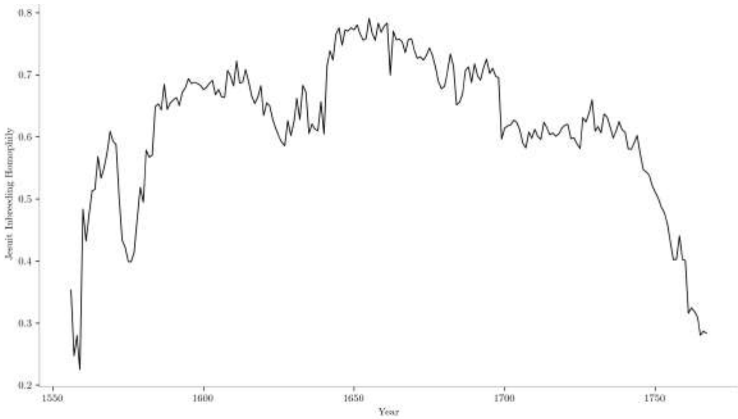


FIGURE 3.9: Jesuit inbreeding homophily over time.

To test this hypothesis, we examine whether Jesuit nodes are densely connected internally while remaining weakly connected to the rest of the network. We define two groups of nodes: Jesuits and non-Jesuits. To quantify Jesuit insularity, we compute the Coleman (1958) inbreeding homophily index, which ranges from -1 (no internal Jesuit connections) to 0 (connections similar to a random network) to 1 (complete inbreeding). Following Currarini, Jackson, and Pin (2009), we define IH as follows:

$$IH_{Jesuits,t} = \frac{H_{Jesuits,t} - \omega_{Jesuits,t}}{1 - \omega_{Jesuits,t}}$$

where $H_{Jesuits,t}$ is the fraction of edges entailing only Jesuits at time t, and $\omega_{Jesuits,t}$ is the relative fraction of Jesuits in the scholar population any given point in time t.

Figure 3.9 plots $IH_{Jesuits,t}$ from 1556 to 1767, the period over which Jesuits were most active. For most of the timeframe, the index remains between 0.6 and 0.8, indicating a high degree of inbreeding. Jesuit universities were typically closed to non-Jesuit professors, and Jesuit scholars rarely taught outside them. However, the index is slightly lower at the beginning, when the Jesuits were establishing their university network, and at the end, preceding their gradual dissolution.[29]

[29]Appendix 3.I.1 provides additional statistics on the Jesuits' position and connectivity in the network, including their number, density, conductance, and decompositions by field.

3.6 Conclusions

We have studied how academic networks in the premodern era influenced the spread of ideas. Using dynamic network models and counterfactual experiments, we showed that features like the emergence of academies and the connections they created across regions helped ideas to spread more widely. By examining the diffusion of groundbreaking ideas and paradigm shifts such as Botanical Realism, Mathematical Astronomy, and Scholasticism, we have validated the role of higher-education institutions in European development.

The counterfactual experiments reveal the nuanced importance of academies: not only did they act as hubs of direct idea dissemination, but they also enhanced the connectivity of the broader network, bridging university communities. Without academies, the spread of ideas born in university settings would have been significantly slower. Our approach also provided insights into regional contributions to scientific progress, highlighting the resilience of the European network of academies and universities, which was dense enough to sustain the circulation of ideas even if certain parts were removed.

While our results are robust to several modelling choices, the analysis is carried out under the assumption that the affiliation network is exogenous, and determined by the data we observe. This is of course a limitation of our approach. We know that the affiliation network results from location choices of individuals, and these choices are determined by several elements, including the rise in prominence of some of its members. Perhaps other scholars joined Tübingen precisely *because* they wanted to learn from Fuchs. Although this endogeneity bias might be of minor importance with respect to ideas themselves, an extension of the analysis could consider exogenous changes in the network, either through shocks to nodes (a popular source of exogenous changes in network is through the unexpected death of some nodes, as in Benzell and Cooke (2021) and Azoulay, Fons-Rosen, and Zivin (2019)), or through direct shocks to edges (as in Cervellati et al. (2025)—improvement of the English postal service—or Chiopris (2024)—delays in building the German train network).

Acknowledgements

The first author acknowledges the support of the European Research Council under the European Union's Horizon 2020 research and innovation program under grant agreement No 883033 "Did elite human capital trigger the rise of the West? Insights from a new database of European scholars." The second author acknowledges the support of the Global PhD Partnership between KU Leuven and UCLouvain on "Bridging Data Science and Intellectual History: Computing the Nodes and Edges in the Old University of Louvain (1425-1797)". The third author acknowledges the support from the Fonds de la Recherche Scientifique-FNRS under Grant n° A2.11903.007-F "Human capital and the rise of the West: the key role of scientific academies".

We are grateful to Jacob Schmutz for his help in dealing with Scholasticism, to Matteo Cervellati for highlighting the interpretation of our network as reflecting an ITT, to Debin Ma for discussing the Chinese Malthus, to Kerstin Enflo for introducing us to Rudbeck and his view on Atlantis, to Sascha Becker, Margherita Fantoli, Violet Soen, Cecilia Garcia Peñalosa, Yann Bramoullé, Catarina Chiopris, Nathan Nunn, Stelios Michalopoulos, Klaus Desmet, Ralf Meizenzhal, Paula Gobbi, Luca Pensieroso, Joachim Voth and to the participants to the Fresh workshop (Louvain-la-Neuve, June 2024) "Institutions, human capital, and long-term development: Lessons from pre-modern Europe", to the XII Héloïse workshop (Warsaw, September 2024) "From East to West and Back: Circulations of Knowledge in Pre-Modern and Modern Europe - Actors, Institutions, and Spaces", to the Guerzensee CEPR workshop, to the conference on Deep-Rooted Factors in Comparative Development (Brown, 2025), and to the Cesifo summer conference in Venice. We also thank the participants to seminars in Lille, Belfast, Naples (Parthenope), Marseille (workshop LORDE 2025), CERGE Prague, and Alice Fabre for her discussion at LORDE 2025.

3.7 Bibliography

Abramitzky, R., & Sin, I. (2014). Book translations as idea flows: The effects of the collapse of communism on the diffusion of knowledge. *Journal of the European Economic Association*, *12*(6), 1453–1520.

Abulafia, A. S. (2011). *Christian–Jewish relations 1000–1300: Jews in the service of medieval christendom*. Pearson.

Ahmadpoor, M., & Jones, B. F. (2017). The dual frontier: Patented inventions and prior scientific advance. *Science*, *357*(6351), 583–587.

Akcigit, U., Caicedo, S., Miguelez, E., Stantcheva, S., & Sterzi, V. (2018). *Dancing with the stars: Innovation through interactions* (Working Paper No. 24466). National Bureau of Economic Research.

Almelhem, A., Iyigun, M., Kennedy, A., & Rubin, J. (2023). Enlightenment ideals and belief in progress in the run-up to the industrial revolution: A textual analysis.

Amburger, E. (1950). *Die mitglieder der Deutschen Akademie der Wissenschaften zu Berlin 1700–1950*. Akademie-Verlag.

Anderson, R. W., Johnson, N. D., & Koyama, M. (2017). Jewish persecutions and weather shocks: 1100–1800. *The Economic Journal*, *127*(602), 924–958.

André, C., & Platteau, J.-P. (1998). Land relations under unbearable stress: Rwanda caught in the Malthusian trap. *Journal of Economic Behavior and Organization*, *34*(1), 1–47.

Applebaum, W. (2003). *Encyclopedia of the Scientific Revolution: From Copernicus to Newton*. Routledge.

Ashraf, Q., & Galor, O. (2011). Dynamics and stagnation in the Malthusian epoch. *American Economic Review*, *101*(5), 2003–41.

Atkin, D., Chen, M. K., & Popov, A. (2022). *The returns to face-to-face interactions: Knowledge spillovers in silicon valley* (tech. rep.). National Bureau of Economic Research.

Azoulay, P., Fons-Rosen, C., & Zivin, J. S. G. (2019). Does science advance one funeral at a time? *American Economic Review*, *109*(8), 2889–2920.

Azoulay, P., Graff Zivin, J. S., & Wang, J. (2010). Superstar extinction. *The Quarterly Journal of Economics*, *125*(2), 549–589.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013). The diffusion of microfinance. *Science*, *341*(6144), 1236498.

Barrett, M. (2023). *The Reformation as Renewal: Retrieving the One, Holy, Catholic, and Apostolic Church*. Zondervan Academic.

Becker, S. O., Lindenthal, V., Mukand, S. W., & Waldinger, F. (2024). Persecution and escape: Professional networks and high-skilled emigration from Nazi Germany. *American Economic Journal: Applied Economics*, *16*(3), 1–43.

Becker, S. O., & Pascali, L. (2019). Religion, division of labor, and conflict: Anti-semitism in Germany over 600 years. *American Economic Review*, *109*(5), 1764–1804.

Bellemare, M. F., & Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, *82*(1), 50–61.

Ben-David, J. (1971). *The scientist's role in society: A comparative study*. Foundations of Modern Sociology Series.

Benzell, S. G., & Cooke, K. (2021). A network of thrones: Kinship and conflict in Europe, 1495–1918. *American Economic Journal: Applied Economics*, *13*(3), 102–33.

Berger-Levrault, O. (1890). *Catalogus professorum Academiarum et Universitatum alsaticarum XVI-XVIII seculi*. Impr. de Berger-Levrault.

Blasutto, F., & De la Croix, D. (2023). Catholic censorship and the demise of knowledge production in early modern Italy. *The Economic Journal*, *133*, 2899–2924.

Borgatti, S. P., & Halgin, D. S. (2011). Analyzing affiliation networks. *The Sage handbook of social network analysis*, 417–433.

Borowiecki, K. J. (2013). Geographic clustering and productivity: An instrumental variable approach for classical composers. *Journal of Urban Economics*, *73*(1), 94–110.

Bramoullé, Y., & Genicot, G. (2024). Diffusion and targeting centrality. *Journal of Economic Theory*, *222*, 105920.

Brunt, L., & García-Peñalosa, C. (2022). Urbanisation and the onset of modern economic growth. *The Economic Journal*, *132*(642), 512–545.

Buringh, E. (2021). The population of European cities from 700 to 2000: Social and economic history. *Research Data Journal for the Humanities and Social Sciences*, *6*(1), 1–18.

Catalán, M. J. (1924). *Historia de la real y pontificia Universidad de Zaragoza*. La Académica.

Cervellati, M., Lazzaroni, S., Marciante, G., & Masella, P. (2025). *The rise of the knowledge economy: Republic of Letters and communication infrastructures in Early Modern England* [University of Bologna].

Chaunu, P. (2014). *Le temps des réformes: La crise de la chrétienté, l'éclatement (1250-1550)*. Fayard.

Chiopris, C. (2024). *The diffusion of ideas* [Harvard University].

Coleman, J. S. (1958). Relational analysis: The study of social organizations with survey methods. *Human organization*, *17*(4), 28–36.

Collège de France. (2018). *Liste des professeurs du Collège de France depuis 1530* [Accessed on December 4, 2024]. https://www.college-de-france.fr/media/en-college/UPL50600_LISTE_DES_PROFESSEURS.pdf

Conrad, E. (1960). *Die Lehrstühle der Universität Tübingen und ihre Inhaber (1477-1927)*.

Copleston, F. C. (1993). *Medieval philosophy: From Augustine to Duns Scotus*. Image Books.

Cross, R. (2024). An accidental Reformation? In *The hanover review: The journal of the london lyceum: 3.1: The reformation as renewal symposium* (pp. 5–13). Hanover Press.

Currarini, S., Jackson, M. O., & Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, *77*(4), 1003–1045.

Curtis, M., & De la Croix, D. (2023). Measuring Human Capital: From WorldCat Identities to VIAF. *Repertorium Eruditorum Totius Europae*, *10*, 17–22.

David, P. A. (1985). Clio and the economics of QWERTY. *The American economic review*, *75*(2), 332–337.

De la Croix, D. (2021). Scholars and literati in European Academia before 1800. *Repertorium Eruditorum Totius Europae*, *5*, 35–41.

De la Croix, D., Docquier, F., Fabre, A., & Stelter, R. (2024). The Academic Market and the Rise of Universities in Medieval and Early Modern Europe (1000-1800). *Journal of the European Economic Association*, *22*, 1541–1589.

De la Croix, D., Doepke, M., & Mokyr, J. (2018). Clans, guilds, and markets: Apprenticeship institutions and growth in the preindustrial economy. *The Quarterly Journal of Economics*, *133*(1), 1–70.

De la Croix, D., & Goñi, M. (2024). Nepotism vs. intergenerational transmission of human capital in academia (1088–1800). *Journal of Economic Growth*.

De la Croix, D., & Morault, P. (2025). Winners and losers from the Protestant reformation: An analysis of the network of European universities. *Journal of Economic History*.

Del Negro, P. (2015). *Clariores: Dizionario biografico dei docenti e degli studenti dell'Università di Padova*. Padova University Press.

Dittmar, J. E. (2011). Information technology and economic change: The impact of the printing press. *The Quarterly Journal of Economics*, *126*(3), 1133–1172.

Domonkos, L. S. (1968). The Polish astronomer Martinus Bylica de Ilkusz in Hungary. *The Polish Review*, 71–79.

Donker, S. (2024). Intelligencers, cliques, and stars in the spread of 17th-century Cartesianism. *Connections*, *44*(1), 4–32.

Drüll, D. (2002). *Heidelberger gelehrtenlexikon: 1386–1651*. Springer.

Dulieu, L. (1979). *La médecine à Montpellier, vol II: La Renaissance*. Les presses universelles.

English, E. D. (2005). *Encyclopedia of the Medieval World*. Facts On File, Inc.

Esperabé de Arteaga, E., et al. (1917). *Historia pragmática e interna de la Universidad de Salamanca*. Imp. y lib. de Francisco Núñez Izquierdo.

Facciolati, J. (1757). *Fasti gymnasii patavini. jacobi facciolati opera collecti ab anno MDXVII quo restitutae scholae sunt ad MDCCLVI*. typis Seminarii.

Ferris, T., Stella, F., & Yon, K. (2010). *The science of liberty: Democracy, reason, and the laws of nature*. Harper New York.

Fogel, R. W. (1964). Railroads and american economic growth: Essays in econometric history.

Fogli, A., & Veldkamp, L. (2021). Germs, social networks, and growth. *The Review of Economic Studies*, *88*(3), 1074–1100.

Frijhoff, W. (1996). Patterns. In H. d. Ridder-Symoens (Ed.), *A history of the University in Europe. vol. ii: Universities in Early Modern Europe (1500–1800)*. Cambridge University Press.

Gabrieli, G. (1989). *Contributi alla storia della accademia dei Lincei*. Accademia nazionale dei Lincei.

Genet, J.-P. (2019). *Projet studium parisiense* [Université Paris 1].

Geraerts, J., & Vasques Filho, D. (2024). Networks of confessional affiliation: Religious choice and the schism of utrecht. *Journal of Historical Network Research*, *10*(1).

Giorcelli, M., Lacetera, N., & Marinoni, A. (2022). How does scientific progress affect cultural changes? a digital text analysis. *Journal of Economic Growth*, *27*, 415–452.

Goyal, S. (2023). *Networks: An economics approach*. MIT Press.

Goyal, S., Van Der Leij, M. J., & Moraga-González, J. L. (2006). Economics: An emerging small world. *Journal of political economy*, *114*(2), 403–412.

Grendler, P. F. (2019). *Jesuit Schools and Universities in Europe 1548–1773*. Brill.

Gunther, R. T. (1925). *Early Science in Oxford, Vol. iv The Philosophical Society*. Oxford.

Gunther, R. T. (1937). *Early Science in Oxford, Vol. xi Oxford colleges and their men of science*. Oxford.

Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in clinical research*, *2*(3), 109–112.

Hallmann, C., Hanlon, W. W., & Rosenberger, L. (2022). Why Britain? the right place (in the technology space) at the right time.

Hanlon, W. W. (2015). Necessity is the mother of invention: Input supplies and directed technical change. *Econometrica*, *83*(1), 67–100.
Hanlon, W. W. (2025). The rise of the engineer: Inventing the professional inventor during the industrial revolution. *The Economic Journal*, ueaf023.
Herbermann, C. G. (1913). *The Catholic Encyclopedia*. Encyclopedia Press, Incorporated.
Howse, D. (1986). The Greenwich list of observatories: A world list of astronomical observatories, instruments and clocks, 1670–1850. *Journal for the History of Astronomy*, *17*(4), i–89.
Jackson, M. O. (2008). *Social and economic networks*. Princeton University Press.
Jedwab, R., Johnson, N. D., & Koyama, M. (2019). Negative shocks and mass persecutions: Evidence from the black death. *Journal of Economic Growth*, *24*(4), 345–395.
Kedzoria, A. (2021). *Zamiowopedia* [Accessed Jan 10 2025]. https://www.zamosciopedia.pl
King, D. (2005). *Finding atlantis: A true story of genius, madness, and an extraordinary quest for a lost world*. Harmony Books.
Koher, A., Lentz, H. H., Hövel, P., & Sokolov, I. M. (2016). Infections on temporal networks—a matrix-based approach. *PloS one*, *11*(4), e0151209.
Koschnick, J. (2025). *Teacher-directed scientific change: The case of the English Scientific Revolution* (tech. rep.). EHES Working Paper, No. 274.
Lamberts, E., & Roegiers, J. (1990). *Leuven university: 1425–1985*. Leuven University Press.
Laouenan, M., Bhargava, P., Eyméoud, J.-B., Gergaud, O., Plique, G., & Wasmer, E. (2022). A cross-verified database of notable people, 3500BC-2018AD. *Scientific Data*, *9*(290).
Luther, M. (1517). *Disputatio contra scholasticam theologiam* [English version "Disputation Against Scholastic Theology" accessed October 14 2024]. https://williamroach.org/2017/08/20/martin-luthers-1517-disputation-against-scholastic-theology/
Maloney, W. F., & Valencia Caicedo, F. (2022). Engineering growth. *Journal of the European Economic Association*, *20*(4), 1554–1594.
Malthus, T. (1807). *An essay on the principle of population* (Fourth Edition). T. Bensley.
Mazzetti, S. (1847). *Repertorio di tutti i professori antichi e moderni della famosa università, e del celebre istituto delle scienze di Bologna*. tipografia di San Tommaso d'Aquino.
McClellan, J. E. (1985). *Science Reorganized: Scientific Societies in the Eighteenth Century*. Columbia University Press.
Meisenzahl, R. R., & Mokyr, J. (2012). The rate and direction of invention in the british industrial revolution. *The rate and direction of inventive activity revisited*, 443–482.
Mokyr, J. (2005). Long-term economic growth and the history of technology. In *Handbook of economic growth* (pp. 1113–1180, Vol. 1). Elsevier.
Mokyr, J. (2011a). *The enlightened economy: Britain and the industrial revolution, 1700-1850*. Penguin UK.
Mokyr, J. (2011b). The gifts of Athena: Historical origins of the knowledge economy. In *The gifts of athena*. Princeton University Press.
Mokyr, J. (2016). *A culture of growth*. Princeton University Press.
Montreal Botanic Garden. (1886). *First annual report 1885*. Gazette Printing Company.
Moser, P. (2005). How do patent laws influence innovation? evidence from nineteenth-century world's fairs. *The American Economic Review*, *95*(4), 1214–1236.
Needham, J. (1964). Science and Society in East and West. *Science and Society*, *28*(4), 385–408. Retrieved July 10, 2025, from http://www.jstor.org/stable/40401068
Ó Gráda, C. (2016). Did science cause the industrial revolution? *Journal of Economic Literature*, *54*(1), 224–239.
Pedersen, O. (1992). *Lovers of learning: A history of the Royal Danish Academy of Sciences and Letters, 1742-1992*. Munksgaard.
Peters, O. (2019). The ergodicity problem in economics. *Nature Physics*, *15*(12), 1216–1221.
Rashdall, H. (1895). *The Universities of Europe in the Middle Ages*. Clarendon Press.
Roller, R. (2023). Tracing the footsteps of ideas: Time-respecting paths reveal key reformers and communication pathways in Protestant letter networks. *SocArXiv*.

Rubin, J. (2014). Printing and Protestants: an empirical test of the role of printing in the Reformation. *Review of Economics and Statistics*, *96*(2), 270–286.

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika*, *69*(1), 239–241.

Schwinges, R. C., & Hesse, C. (2019). *Repertorium Academicum Germanicum* [Consulted October 2, 2024]. https://en.rag-online.org/

Serafinelli, M., & Tabellini, G. (2022). Creativity over time and space: A historical analysis of European cities. *Journal of Economic Growth*, *27*(1), 1–43.

Silberman, L. (1960). Hung Liang-Chi: A Chinese Malthus. *Population Studies*, *13*(3), 257–265.

Slottved, E. (1978). *Lærestole og lærere ved københavns universitet 1537-1977*. Samfundet for dansk Genealogi og Personalhistorie.

Smith, J., Anthony A. (2008). Indirect inference. In S. N. Durlauf & L. E. Blume (Eds.), *The new palgrave dictionary of economics*. Palgrave Macmillan.

Stelter, R., De la Croix, D., & Myrskylä, M. (2021). Leaders and Laggards in Life Expectancy among European Scholars from the Sixteenth to the Early Twentieth Century. *Demography*, *58*, 111–135.

The British Library. (2021). Database of Italian academies [Consulted on June 6, 2023]. https://www.bl.uk/catalogues/ItalianAcademies/About.aspx

The Editors of Encyclopaedia Britannica. (2024). Guy de Chauliac [Accessed: 27 January 2025]. https://www.britannica.com/biography/Guy-de-Chauliac

Udías, A. (2003). *Searching the heavens and the earth: The history of jesuit observatories* (Vol. 286). Springer Science & Business Media.

Uniwersytet Jagielloński. (2019). *Corpus Academicum Cracoviense* [Accessed: January 20, 2021]. http://www.archiwum.uj.edu.pl/corpus-academicum-cracoviense1

Vidal y Díaz, A., et al. (1869). *Memoria histórica de la Universidad de Salamanca*. Imprenta de Olivo y Hermano.

Voigtländer, N., & Voth, H.-J. (2012). Persecution perpetuated: The medieval origins of anti-semitic violence in Nazi Germany. *The Quarterly Journal of Economics*, *127*(3), 1339–1392.

Von Bahr, G. (1945). *Medicinska fakulteten i Uppsala*. Almqvist & Wiksell International.

von Aschbach, J. R. (1865). *Geschichte der wiener universität und ihre gelehrten*. Verlag der Universität.

Wootton, D. (2015). *The invention of science: A new history of the scientific revolution*. Penguin UK.

Xue, M. (2025). *Enlightenment under autocracy: The origins of liberalism in China* [LSE].

Zamani, M., El-Hajj, H., Vogl, M., Kantz, H., & Valleriani, M. (2023). A mathematical model for the process of accumulation of scientific knowledge in the Early Modern period. *Humanities and Social Sciences Communications*, *10*(1), 1–10.

Zanardello, C. (2024). *Early modern academies and growth* [LIDAM Discussion Paper - 2024/12].

Zenou, Y. (2016). Key players. *The Oxford Handbook of the Economics of Networks*, *4*, 244–274.

Zhao, D., & Strotmann, A. (2015). *Analysis and visualization of citation networks*. Morgan & Claypool Publishers.

3.A Descriptive Statistics

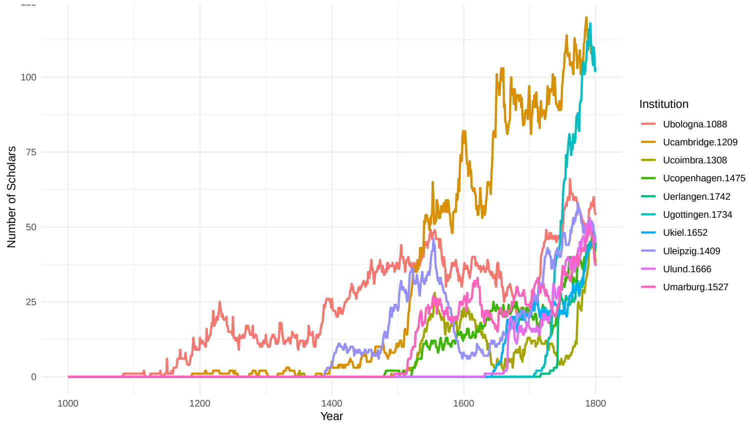


FIGURE 3.10: Number of university professors between 1000 and 1800 of the top 10 universities considering the number of scholars in 1793. In the legend, the name of the institution with the foundation date.

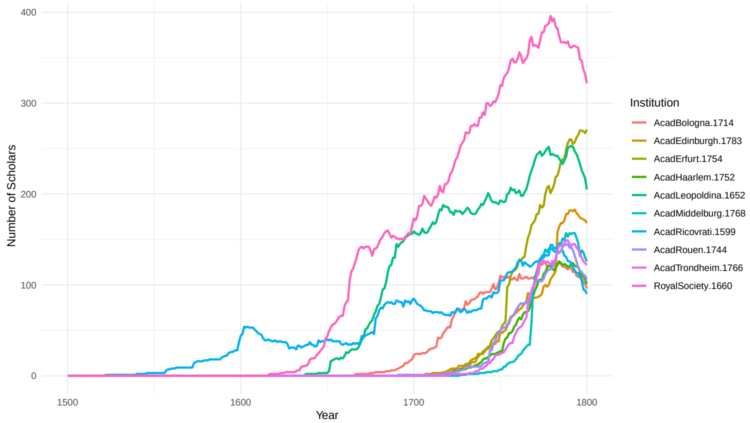


FIGURE 3.11: Number of academicians between 1500 and 1800 of the top 10 academies considering the number of scholars in 1793. In the legend, the name of the institution with the foundation date.

Institution	Mean	Med	SD	Min	Q1	Q3	Max
Ubologna.1088	34.14	36	14.39	13	25	43	54
Ucambridge.1209	59.29	93	50.36	1	8	100	105
Ucoimbra.1308	8.57	4	13.78	0	0	9	38
Uerlangen.1742	9.71	0	17.39	0	0	12.5	43
Ugottingen.1734	23.00	0	41.37	0	0	29	103
Ukiel.1652	12.71	4	16.42	0	0	21	43
Ucopenhagen.1475	17.71	22	17.58	0	1	27.5	45
Uleipzig.1409	21.43	20	19.17	0	7	34	48
Ulund.1666	14.00	1	19.26	0	0	25.5	46
Umarburg.1527	16.43	17	16.67	0	0.5	30	37

TABLE 3.9: Descriptive statistics for the top 10 universities, considering the number of scholars in 1793. In the first column, the name of the institution with the foundation date. “Med” stands for Median, “SD” for Standard Deviation, “Min” for minimum, “Q1” for first quantile, “Q3” for third quantile, and “Max” for maximum.

Institution	Mean	Med	SD	Min	Q1	Q3	Max
ABologna.1714	32.71	0	48.91	0	0	59.5	110
AEdinburgh.1783	30.86	0	63.06	0	0	24	168
AErfurt.1754	45.71	0	101.01	0	0	24.5	271
AHaarlem.1752	18.29	0	37.84	0	0	13.5	101
ALeopoldina.1652	80.00	3	99.80	0	0	176	205
RoyalSociety.1660	122.57	44	148.71	0	0	246.5	322
ARicovrati.1599	45.71	40	47.12	0	0	87.5	105
ARouen.1744	22.43	0	41.26	0	0	25.5	106
ATrondheim.1766	20.86	0	45.49	0	0	12	122
AMiddelburg.1768	19.00	0	47.25	0	0	3.5	126

TABLE 3.10: Descriptive statistics for the top 10 academies, considering the number of scholars in 1793. In the first column, the name of the institution with the foundation date. “Med” stands for Median, “SD” for Standard Deviation, “Min” for minimum, “Q1” for first quantile, “Q3” for third quantile, and “Max” for maximum.

3.B Contextual information

3.B.1 Botanical Realism and botanic gardens

In Europe, natural history traces its roots back to ancient Greek philosophers such as Aristotle, Theophrastus, and Dioscorides. During the Scientific Revolution, botany underwent major advancements, transitioning from a primarily descriptive field into a more systematic and experimental science. By the 16^{th} and 17^{th} centuries, botany started encompassing not only the identification and classification of plants species, but also the growing field of plant physiology, investigating the properties and functions of plants life. This marked a shift in botanical practices, expanding beyond the descriptive and illustrative focus of ancient authors (Applebaum, 2003), towards a more empirical approach that we will refer to as “Botanical Realism”. Before 1650, botany was considered merely a complement to medical studies, but it became an independent field of study during the Scientific Revolution. Universities renown for their medical faculties began offering innovative botany lectures, where students were taken directly to gardens to observe plant species first-hand. These universities were also the first to establish their own botanic gardens to support further research and development in botany. Following this trend, private citizens and local lords also recognized the importance of botanical studies and funded the creation of such gardens (Applebaum, 2003).

One key figure in this transformation was Leonhart Fuchs, a German physician and botanist. He is best known for his book *De historia stirpium commentarii insignes*, which translates to “Notable commentaries on the history of plants". First printed in Basel in 1542, one year before Nicolaus Copernicus’ *De revolutionibus orbium coelestium* and Andreas Vesalius’ *De humani corporis fabrica*, this work laid the foundation for modern botany. Fuchs not only provided ideal visual representations of 511 plant species, but he also included his own critical observations on their uses and characteristics, highlighting differences from ancient texts (Applebaum, 2003). Figure 3.12 shows a page of this book, emphasizing the realistic description of a plant.

Furthermore, we gathered information on the existence and founding dates of European botanic gardens. Our starting point was the first annual report of the Montreal Botanic Garden (1886), which lists the botanic gardens open worldwide in 1885. From this, we selected only European gardens and determined their founding dates using AI-assisted tools, which were then manually verified through sample checks. We then matched this sample of botanic gardens with our university cities, assuming that a city without a botanic garden was not listed in the first annual report by Montreal Botanic Garden (1886). To fix ideas, Figure 3.13 shows the *Hortus botanicus* (botanic garden) of the University of Leiden, which was opened in 1590. This is a case of a large garden with a

FIGURE 3.12: A page of *De historia stirpium commentarii insignes*.

dedicated building, but some other gardens were much smaller.

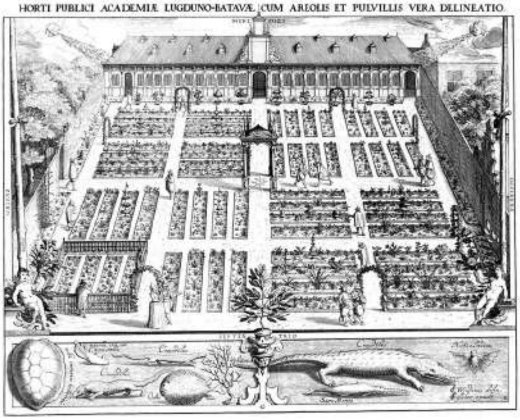


FIGURE 3.13: The Hortus botanicus of Leiden.

3.B.2 Mathematical Astronomy and astronomical observatories

The 15th and 16th centuries witnessed a growing interest in experimental science and increasing dissatisfaction with the explanations offered by ancient astronomical authorities, such as Claudius Ptolemy (c. 100–c.170, Alexandria). Similar to the approach seen in Botanical Realism, this paradigm shift led mathematicians and astronomers to question the accuracy of Ptolemy's models and refine them through observation and mathematical analysis. This era marked the beginning of the astronomical revolution, characterized by advances in trigonometry, new geometric formulas, and the adoption of decimal calculations in astronomy. The focus shifted from simply explaining celestial motions to understanding the physical mechanisms behind them.

A key figure in this revolution was Regiomontanus, pseudonym of Johannes Müller. His mastery of Greek and mathematics enabled him to study the original works of Ptolemy and other ancient thinkers. At the University of Vienna, around 1454, he and his mentor, Georg Peurbach (1423–1461), began collaborating on *Theoricæ novæ planetarum*. This seminal work introduced new methods for solving plane and spherical trigonometry problems, including the use of sine and tangent functions. Regiomontanus also created extensive trigonometric tables with values calculated to decimal units, which remained influential for centuries. As such, he can be considered a pioneer of Mathematical Astronomy. While trigonometry had been used in astronomy and other sciences, Regiomontanus's contributions greatly enhanced its application. His work, alongside Peurbach's, laid the groundwork for later revolutionary astronomers such as Copernicus, Kepler, and Galileo (Applebaum, 2003).

After Peurbach's death, Regiomontanus moved to Northern Italy, then to Hungary in 1467. Later, he settled in Nuremberg, drawn by its status as a free city and central location. There, he established a workshop and printing press, dedicating himself to the dissemination of scientific knowledge. In 1463, he published *Epitoma in Almagestum Ptolemaei*, which clarified, corrected, and expanded Ptolemaic astronomy. Two pages of *Epitoma* are shown if Figure 3.14. In 1472, he published *Theoricæ novæ planetarum*, his collaboration with Peurbach. In 1475, Pope Sixtus IV invited him to Rome to work on calendar reform, but Regiomontanus died shortly after, at age 41.

We examine the creation of observatories in the same 185 university cities defined in Section 3.3.2. We collected the names and foundation dates of observatories from *The Greenwich List of Observatories* compiled by Howse (1986). As with the botanic gardens, we only considered observatories in continental Europe, assuming that a location lacked one if it was not listed in our source. To fix ideas, Figure 3.15 shows the main building of the University of Prague, the Clementinum. There, a tower was built in 1722. Later, in 1751, instruments were installed, and the tower became an astronomical observatory.

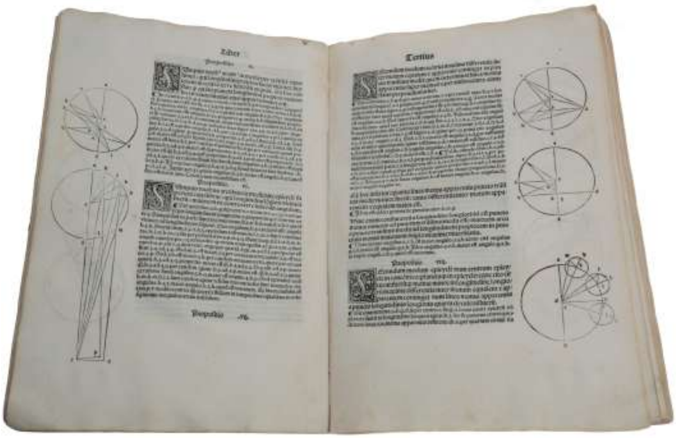


FIGURE 3.14: Two pages of *Epitoma in Almagestum Ptolemaei*.

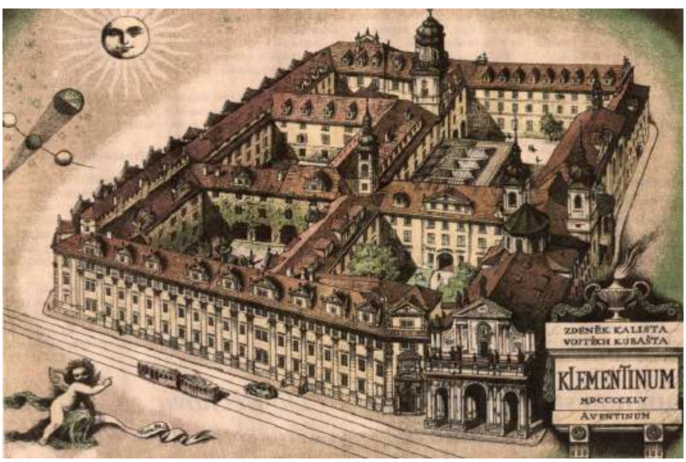


FIGURE 3.15: The *Clementinum* in Prague with its observatory.

3.C Empirical Assessments

3.C.1 Descriptive Statistics

Table presents some descriptive statistics, which supports the results in the main text, Table 3.2 and Table 3.3.

TABLE 3.11: Summary statistics.

Variable	NAs	Obs	Mean	Med	SD	Min	Max
(ihs) Exposure to Botanical Realism	0	54390	0.31	0	0.83	0	5.75
(ihs) Non exposure to Botanical Realism	0	54390	0.57	0	1.19	0	6.14
(ihs) Exposure to Math. Astronomy	0	54390	0.50	0	1.25	0	7.32
(ihs) Non exposure to Math. Astronomy	0	54390	0.42	0	0.99	0	5.68
(ihs) City pop. in 1500	4410	49980	2.80	2.78	1.06	0	5.52
(ihs) Dist. to Tübingen	0	54390	7.04	7.15	0.87	0	8.34
(ihs) Dist. to Vienna	0	54390	7.38	7.44	0.81	0	8.44

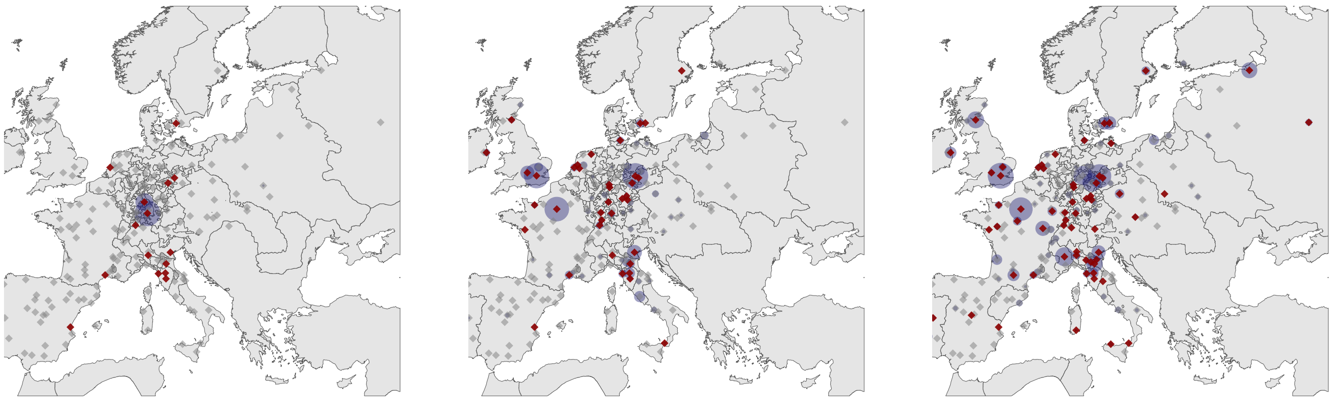
(ihs) refers to the transformation in inverse hyperbolic sine of the relative variable. The 4410 missing values for city population indicates that for 15 cities (and 294 years) we do not have population data. "NAs" stands for missing values, "Obs" for observation counts, "Med" for Median, "SD" for Standard Deviation, "Min" for minimum, and "Max" for maximum.

3.C.2 Botanical Realism

Figure 3.16 shows the yearly exposure to Botanical Realism in three different points in time. This figure visualizes how exposure evolves over time, allowing us to trace how these innovative ideas were initially concentrated around Tübingen in 1600, spread across Europe by 1700, and further expanded, reaching smaller and more distant urban centers by 1793. Blue bubbles represent exposure to Botanical Realism, while red diamonds indicate cities that had at least one botanic garden by the respective year.

3.C.3 Mathematical Astronomy

Figure 3.18 visualizes the sample of universities cities with their yearly exposure to Mathematical Astronomy in 1600, 1700, and 1793. This figure allows to



(A) Year 1600 (B) Year 1700 (C) Year 1800

FIGURE 3.16: Blue bubbles represent the yearly exposure to Botanical Realism in years 1600, 1700, and 1793, respectively. $\alpha = 0.25$ and $D = 5{,}000$. Cities with a botanic garden are shown as red diamonds; others are grey diamonds.

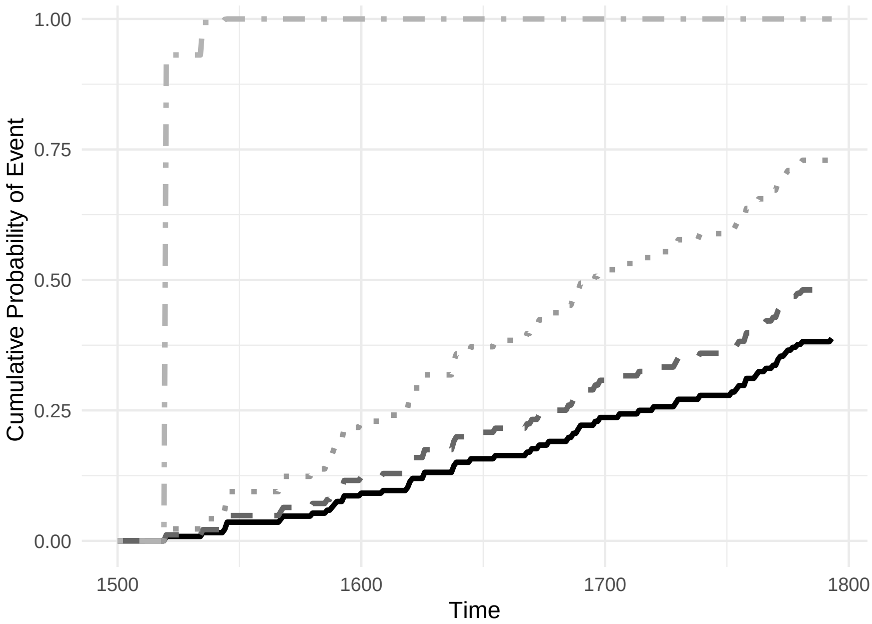
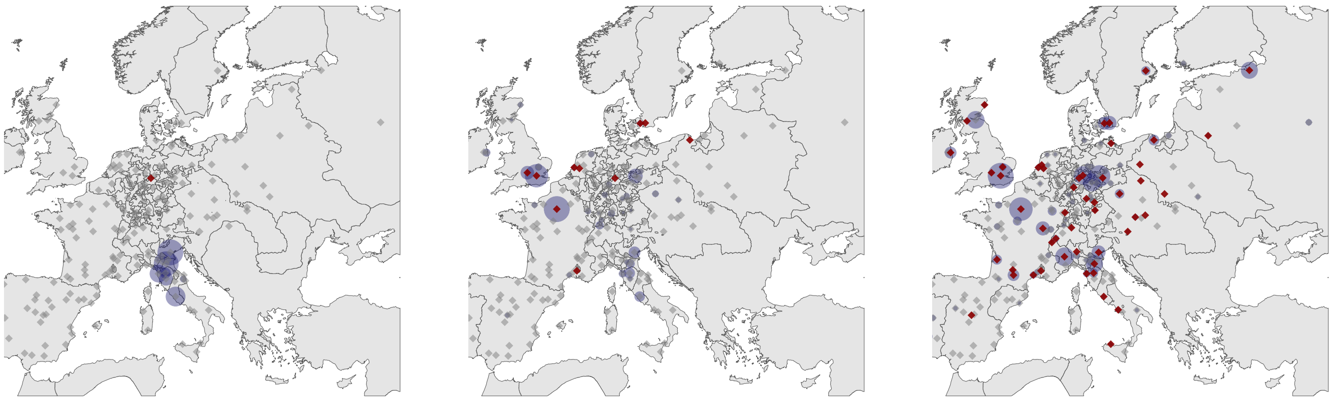


FIGURE 3.17: Probability of getting a botanic garden by time for different exposure to Botanical Realism levels: dot-dashed line considers a constant exposure of 5.75 (max exposure), the dotted line a constant exposure of 1, the dashed line a constant exposure of 0.31 (mean exposure), and the solid line a constant null exposure.

see how the exposure evolves over time and its interaction with the creation of observatories. As in the first experiment, blue bubbles represent exposure to Mathematical Astronomy, while red diamonds indicate cities that had at least one observatory in the relative year.



(A) Year 1600 (B) Year 1700 (C) Year 1800

FIGURE 3.18: Blue bubbles represent the yearly exposure to Mathematical Astronomy in years 1600, 1700, and 1793, respectively. $\alpha = 0.25$ and $D = 5{,}000$. University cities with an astronomical observatory are the red diamonds; cities without are grey diamonds.

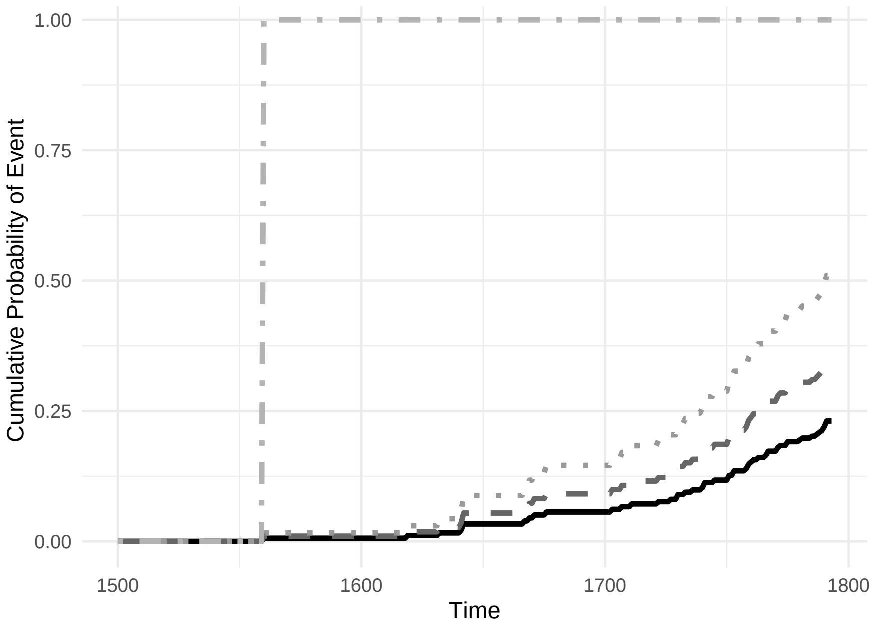
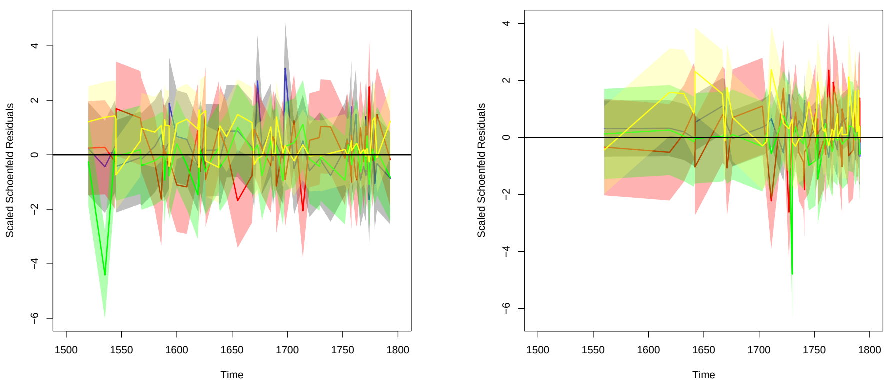


FIGURE 3.19: Probability of getting an Observatory by time for different exposure to Mathematical Astronomy levels: dot-dashed line considers a constant exposure of 7.32 (max exposure), the dotted line a constant exposure of 1, the dashed line a constant exposure of 0.5 (mean exposure), and the solid line a constant null exposure.

3.D Tests for the proportionality of Hazard Functions

In this section we test the proportionality of the hazard functions (e.g.,scaled Schoenfeld residuals) of all the covariates of Table 3.2 and Table 3.3, respectively, and time. The global correlation is only slightly significant in the case of Botanical Realism, while it is not at all significant for Mathematical Astronomy. We can see that by looking at the confidence intervals almost always overlapping with the zero line. The blue line and grey confidence interval relates to the hazard ratios of "(ihs) Exposure to Botanical Realism" (Fig 3.20a) and "(ihs) Exposure to Mathematical Astronomy" (Fig 3.20b), respectively. The yellow line and confidence interval correspond to the hazard ratios of "(ihs) Non exposure" to Botanical Realism (Panel a) and to Mathematical Astronomy (Panel b). The red line and confidence interval correspond to the scaled Schoenfeld residuals of "(ihs) City population in 1500", while the green line and interval of confidence represent the hazard ratios of "(ihs) Distance to Tübingen" (Fig 3.20a) and "(ihs) Distance to Vienna" (Fig 3.20b), respectively.



(A) Botanical Realism

(B) Mathematical Astronomy

FIGURE 3.20: Joint correlations between the scaled Schoenfeld residuals (e.g., hazard ratios) of all the covariates in column (3) of (a) Table 3.2 and time, and (b) Table 3.3 and time.

3.E Example of scholastic reasoning

To fix ideas, a good example of reasoning using the tools of scholastic theology is the second proof of the existence of God by Aquinas as reported by Copleston (1993). Remark that it does not rely much on the scriptures, but rather on a kind of mathematical/logical argumentation:

1. In the world, we can see that things are caused.
2. But it is not possible for something to be the cause of itself because this would entail that it exists prior to itself, which is a contradiction.
3. If that by which it is caused is itself caused, then it too must have a cause.
4. But this cannot be an infinitely long chain, so, there must be a cause which is not itself caused by anything further.
5. This everyone understands to be God.

3.F Augmenting Rubin's data

Rubin (2014) compiled data on whether European cities were Protestant in 1530, 1560, and 1600, focusing primarily on cities within the Holy Roman Empire. His approach assumed that cities located in officially Catholic countries remained Catholic throughout the period. However, this assumption does not hold for regions such as France and the Low Countries (modern-day Belgium), where several cities adopted Protestantism temporarily before being reconquered by Catholic forces.

To address this limitation, we have updated the religious status of the following cities to reflect periods of Protestant control, based on information from the Catholic Encyclopedia (Herbermann, 1913).

- Die. Protestant control: 1562-1628/1629. Became a Protestant stronghold in Dauphiné during the Wars of Religion. Occupied by royal troops during Richelieu's repression of Huguenot fortresses.
- La Rochelle. Protestant period: 1550s-1628. Key Huguenot stronghold. Besieged and subdued by royal forces under Richelieu in 1628.
- Montauban. De facto Protestant Rule: 1561-1629. Montauban became one of the most fortified and independent Huguenot cities in France. Finally capitulated (1629) to royal troops under Richelieu after the fall of La Rochelle and a renewed campaign to suppress Huguenot political autonomy.

- Montpellier. Protestant control: 1562-1622. Served briefly as a de facto capital of Huguenot political assemblies. Lost military and political autonomy after Siege of Montpellier (1622) by royal troops under Louis XIII.
- Nîmes. Protestant control: 1561-1629. After violent iconoclasm in 1561, the city came under Huguenot control. Held by Protestants throughout the Wars of Religion; formally lost autonomy in 1629 after Richelieu's campaigns.
- Ostend. In 1572, Ostend joined the Dutch Revolt and came under the control of the rebel States-General of the Netherlands, aligning with Protestant (Reformed) forces. After the Siege of Ostend (1601-1604), the city was almost entirely destroyed, and Catholicism was restored under Spanish Habsburg rule.
- Uzès. Protestant control: ca. 1562-1629. Important center in the Languedoc region. Remained predominantly Protestant until submission to royal forces in 1629, when Richelieu dismantled Protestant strongholds.

For the sake of completeness, we list here the cities with Protestant control outside Rubin's years 1530,1560, and 1600.

- Antwerp. Protestant period: ca. 1566-1585. During the Reformation, Antwerp became a Calvinist stronghold. After the Fall of Antwerp in 1585, Protestant worship was banned, and many Protestants fled north.
- Caen. Protestant control: 1562-1572. Important Protestant stronghold in Normandy. The St. Bartholomew's Day Massacre (1572) led to widespread killings and ended Protestant rule.
- Ghent. Protestant period: ca. 1577-1584. Calvinist Republic of Ghent established during the Dutch Revolt. Ended with Spanish reconquest.
- Tournai. Protestant control: 1577-1581. Strong Calvinist presence; fell to Spanish forces in 1581.

3.G Scholasticism: Additional Results

In Table 3.12 we show the same main explanatory variables as in Table 3.4 in the main text but we include more controls and fixed effects. The results are robust. In addition to controlling for the presence of all universities active in 1500 as indicated in our database (De la Croix, 2021), we also include the city populations in 1500 taken from Buringh (2021), transformed in inverse hyperbolic sine (ihs) to account for cities with no recorded population estimates.[30] The remaining control variables are selected from Rubin (2014) and include factors related to the economic status of cities: the presence of a printing press by 1500, whether the city was a free city by 1517, the market potential of the city in 1500, its membership in the Hanseatic League by 1517, whether it hosted a bishop or archbishop by 1517, and whether the city had direct access to water. We also include Fixed Effects capturing time-invariant characteristics common to each imperial circle and historical country as in 1500.

[30] The inverse hyperbolic sine is similar to the logarithmic transformation but can accommodate zero (Bellemare & Wichman, 2020), which makes it particularly useful when dealing with the large number of cities from Rubin (2014) with no available population estimates in Buringh (2021).

TABLE 3.12: Linear Probability Model - Exposure to Scholasticism in 1508 and cities' probability to become protestant in 1530, 1560, and 1600.

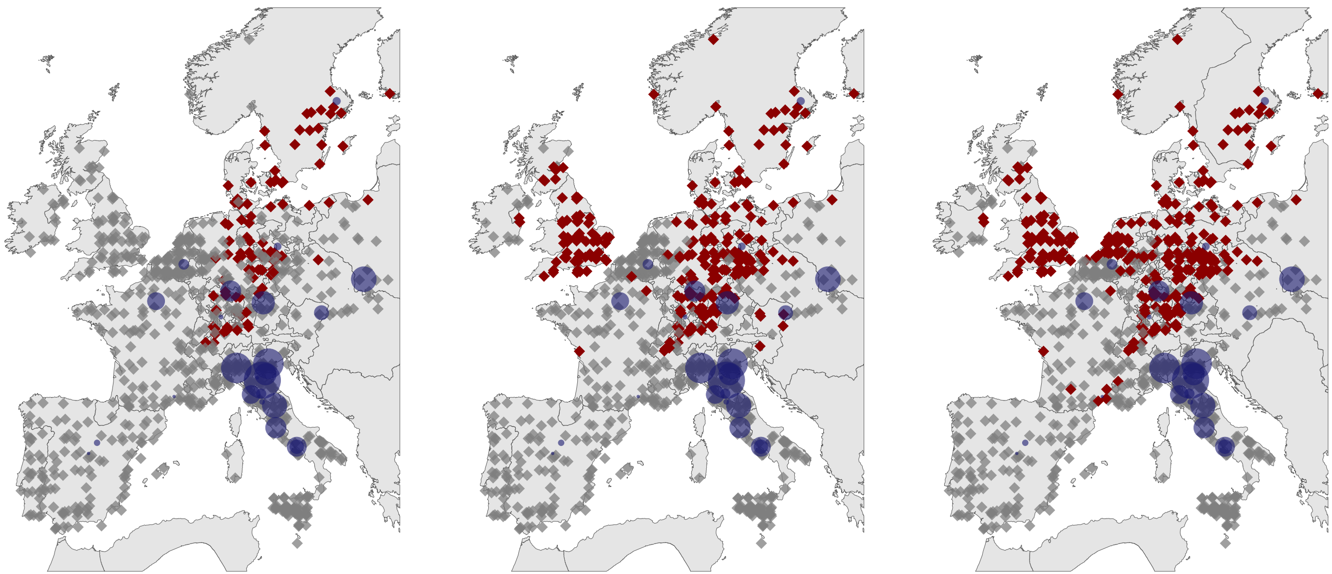
	Protestant in			Protestant in		
	1530 (1)	1560 (2)	1600 (3)	1530 (4)	1560 (5)	1600 (6)
Exposure Scholasticism S^c_{1508}	0.001*	0.003***	0.003***	0.001*	0.003***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Non exposure to Scholasticism $\check{S}^c_{1508}$				-0.002	-0.007	-0.001
				(0.003)	(0.005)	(0.007)
Presence of university in 1500	0.003	-0.007	-0.006	0.004	-0.003	-0.006
	(0.021)	(0.030)	(0.031)	(0.022)	(0.031)	(0.031)
Printing press by 1500	-0.043*	-0.051**	-0.059***	-0.043*	-0.048**	-0.060***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.022)
(ihs) City population in 1500	0.012**	0.005	0.005	0.012**	0.005	0.005
	(0.005)	(0.007)	(0.007)	(0.005)	(0.007)	(0.007)
Free Imperial City by 1517	0.120	0.179*	0.274***	0.119	0.174*	0.274***
	(0.082)	(0.098)	(0.104)	(0.082)	(0.097)	(0.104)
Market potential in 1500	-0.006**	-0.014**	-0.013**	-0.006**	-0.013**	-0.013**
	(0.003)	(0.006)	(0.006)	(0.003)	(0.006)	(0.006)
Hanseatic by 1517	0.024	0.080	0.082*	0.024	0.081	0.082*
	(0.038)	(0.052)	(0.049)	(0.039)	(0.052)	(0.049)
Lay magnate	-0.014	0.149**	0.168**	-0.015	0.146**	0.169**
	(0.038)	(0.067)	(0.071)	(0.038)	(0.069)	(0.071)
(Arch)Bishop by 1517	-0.035*	-0.057**	-0.062***	-0.033*	-0.052**	-0.063**
	(0.019)	(0.025)	(0.024)	(0.019)	(0.025)	(0.025)
Access to water	0.008	-0.0003	-0.004	0.008	-0.001	-0.005
	(0.016)	(0.019)	(0.019)	(0.016)	(0.019)	(0.020)
Imperial Circle FE	YES	YES	YES	YES	YES	YES
1500 Country FE	YES	YES	YES	YES	YES	YES
Observations	867	867	867	867	867	867
Adjusted R^2	0.501	0.715	0.733	0.501	0.715	0.733
Log Likelihood	110.49	27.35	15.84	110.62	28.65	15.86

Notes: Robust SE clustered by territory in parentheses. A constant term is included in all regressions.
Dependent variable 'Protestant' takes value 1 if the city is protestant in 1530, 1560, 1600, respectively. Data on cities' religion taken from Rubin (2014) and updated as in Appendix 3.C.3.
"Presence of university in 1500" is a dummy variable taking value 1 if the city had a university in 1500 as in our database (De la Croix, 2021). "Exposure to Scholasticism S^c_{1508}" and "Non Exposure to Scholasticism $\check{S}^k_t$" are computed as in equations 3.4 and 3.7, respectively. The remaining control variables are selected from Rubin (2014).

3.H Additional Results with Practical Surgery

In addition to the results in Table 3.4 and Table 3.12, where we include 'Non exposure to scholasticism' to control for other orthogonal pathways that might have influenced the shift towards Protestantism, we present results controlling for a specific, orthogonal idea: *Exposure to Practical Surgery* PS^c_{1508}. This is interesting because it captures a distinct effect, separate from Scholasticism. We compute exposure to Practical Surgery similarly to Scholasticism, with the only difference being the starting point. Its "inventor", Guy de Chauliac (c. 1300–1368), was the most prominent physician and surgeon of the Middle Ages. His most famous work *Chirurgia Magna*, published in 1363, was the first to detail surgical procedures, previously handled mostly by charlatans. It remained the main reference well into the 17$^{\text{th}}$ century (The Editors of Encyclopaedia Britannica, 2024).

The orthogonal relationship between Scholasticism and Practical Surgery is evident in Figure 3.21 (Appendix 3.C.3): Practical Surgery was more prominent in Southern Europe, spreading independently of Scholasticism. In the linear probability model shown in Columns (1)–(3) of Table 3.13, *Exposure to Practical Surgery* PS^c_{1508} always has a negative sign and is significant in 1560 and 1600. We interpret this as further evidence of robustness: even when controlling for university presence and exposure to an orthogonal intellectual tradition, the correlation between Scholasticism and a city's likelihood of becoming Protestant remains positive and significant. This reinforces our hypothesis of a "disgust" effect triggered by Scholasticism.



(A) Protestant cities in 1530 (B) Protestant cities in 1560 (C) Protestant cities in 1600

FIGURE 3.21: Blue bubbles represent the exposure to Practical Surgery 30 years prior to 1508, $\alpha = 0.25$ and $D = 5{,}000$. Protestant cities are the red diamonds, and Catholic cities are the grey diamonds. Data on cities' religion are taken from Rubin (2014) and updated as in Appendix 3.C.3.

TABLE 3.13: Linear Probability Model - Exposure to Scholasticism in 1508 and cities' probability to become protestant in 1530, 1560, and 1600.

	Protestant in		
	1530	1560	1600
	(1)	(2)	(3)
Exposure to Scholasticism S^c_{1508}	0.001**	0.004***	0.006***
	(0.001)	(0.001)	(0.001)
Presence of university in 1500	-0.027	-0.042	-0.082*
	(0.027)	(0.044)	(0.045)
Exposure to Practical Surgery PS^c_{1508}	-0.001	-0.003**	-0.005***
	(0.001)	(0.001)	(0.001)
Observations	867	867	867
Adjusted R^2	0.022	0.139	0.259
Log Likelihood	-197.79	-467.50	-443.61

Notes: Robust SE clustered by territory in parentheses. A constant term is included in all regressions.

3.I Placebo Inventors of Botanical Realism

Here, we list the twelve scholars we use as "placebo" inventors in the counterfactual analysis in Section 3.5.1. While it is speculative to say whether each of these twelve individuals could have invented Botanical Realism, many of them were indeed prominent scholars in fields that could have contributed to the development of a more empirical approach to botany. However, the emergence of a paradigm like Botanical Realism depended on a combination of factors–intellectual, cultural, and scientific–beyond the work of individual scholars. Below, a closer look at the potential of each of the individuals:

- Juan Aguilera was professor of medicine and sciences at the University of Salamanca from 1538 to 1560 (Esperabé de Arteaga et al., 1917; Vidal y Díaz et al., 1869): as he had a background in natural philosophy or medicine, he could have contributed to a more empirical study of plants, as Salamanca was a leading university with a strong focus on scientific inquiry during the Renaissance.
- Gaspard Lax de Sarenina was professor of sciences at the University of Zaragoza from 1521 to 1560 (Catalán, 1924): Known for his work in mathematics and philosophy, Lax might not have had direct expertise in

botany, but scholars in these fields often contributed to broader scientific shifts.

- John Warner taught medicine at the University of Oxford from 1520 to 1554 (Gunther, 1937) and was a member of the Royal College of Physicians (1561): he might have had access to Renaissance humanist ideas, but Oxford was more conservative at the time, and Warner would need a strong inclination toward natural science to spearhead Botanical Realism.
- Jeremius Dryvere taught medicine at the University of Louvain from 1522 to 1554 (Lamberts & Roegiers, 1990): Louvain was a center of scientific learning, and someone like Dryvere could have contributed to botanical studies.
- Andreas Goldschmidt taught medicine at the University of Königsberg from 1550 to 1559 (Schwinges & Hesse, 2019): As a scholar trained in Wittenberg, where humanism and scientific inquiry were encouraged, Goldschmidt could have been part of the intellectual currents that led to developments like Botanical Realism.
- Mikołaj Mleczko Wieliczki was professor of medicine at the University of Cracow from 1512 to 1552 (Uniwersytet Jagielloński, 2019): Cracow had a strong tradition in astronomy and natural sciences, and a scholar like Wieliczki could have contributed to the empirical study of nature.
- Jacob Bording was professor of medicine at the University of Rostock from 1549 to 1556 and at the University of Copenhagen from 1556 to 1560 (Slottved, 1978): as a prominent physician, Bording would likely have been interested in botany as it related to medicine, which was a key motivator for many early botanists.
- Antoine Saporta was professor of medicine at the University of Montpellier from 1531 to 1573 (Dulieu, 1979): Montpellier was a leading medical school, and Saporta, as a physician, would have had a strong interest in medicinal plants. He could have been well-positioned to develop a more scientific approach to botany.
- Girolamo Donzellini taught medicine at the University of Padua from 1541 to 1543 (Facciolati, 1757): The University of Padua was a hub of medical and scientific learning, so Donzellini, with his interest in medicine, might have had the right environment to develop Botanical Realism.
- Oronce Fine taught sciences at the Royal College in Paris from 1530 to 1555 (Collège de France, 2018) : Although primarily a mathematician and cartographer, Fine was part of a broader Renaissance movement that emphasized empirical study, and he could have contributed to a more systematic approach to botany.
- Realdo Colombo taught medicine at the universities of Padua (1538–1544), Pisa (1544–1548) and Roma (1548–1559), see Del Negro (2015): he was

a noted anatomist, and his empirical methods in anatomy could have translated well into botany, particularly in the detailed study of plant structures.

- Georg Joachim Porris taught sciences at the universities of Wittenberg (1537–1542), Leipzig (1542–1551) and Vienna (1554–1555), see Schwinges and Hesse (2019) and von Aschbach (1865). Also known as Rheticus, Porris was an astronomer and mathematician. While not a botanist, his scientific mindset might have inclined him toward an empirical approach in natural studies if he had turned his attention to plants.

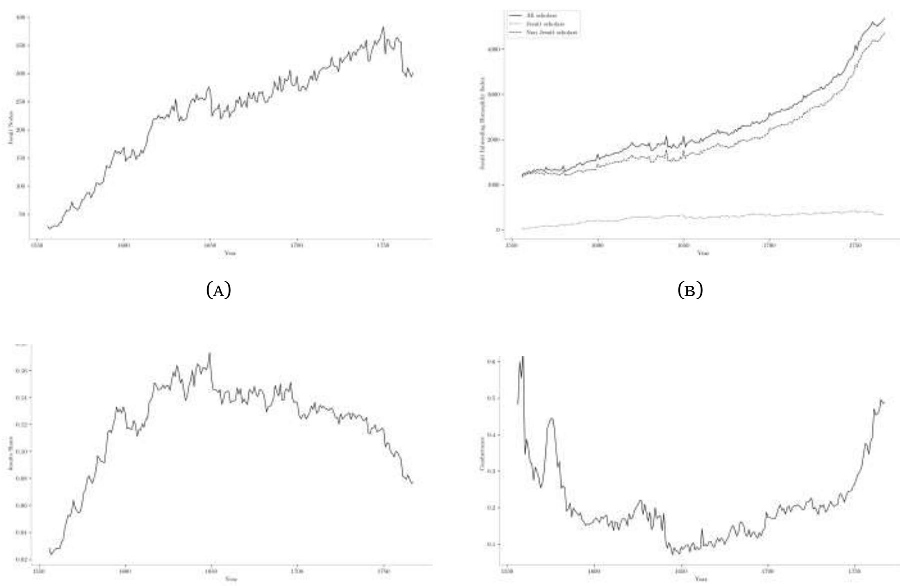
3.I.1 Centrality of the Placebo inventors

The likelihood of an idea spreading depends on how well-connected its inventor is, typically measured by degree centrality—the number of edges a node has. However, degree centrality (and other centrality measures) does not fully capture the dynamics of idea diffusion, which unfolds over time rather than at a single moment. A high degree centrality does not necessarily equate to greater reach. To illustrate this, we report the degree centrality of each placebo inventor in 1542 (i.e., the number of colleagues in the same field) and in 1550, when the inventor was unconnected in 1542.

If Fine (2) or Colombo (19) had been the originators, the idea would have already spread to 50% of the academic population by the second half of 1600. The next fastest spread would have occurred with Dryvere (8) and Porris (22), followed by Saporta (6) and Bording (2 in 1550). For Warner (10), the spread would occur significantly later, only accelerating after the establishment of the first major academies, with a sharp increase around 1660. If Goldschmidt (0, 3 in 1550) had been the inventor, the idea would have struggled to survive initially, only gaining rapid traction around 1680. Interestingly, had Aguilera (8) been the inventor, the idea would have remained confined to Salamanca, persisting without spreading elsewhere until the end of the 18th century, when we observe a sudden spike. This shift coincides with scientists from the Spanish university beginning to affiliate with more international academies. In two cases (Wieliczki (3) in Cracow, and Lax (0) in Zaragoza), the idea fails to spread. These simulations demonstrate how academic institutions can play a crucial role in preserving ideas that might otherwise remain obscure due to their development in less influential locations.

3.J Placebo networks and the role of the Jesuits

In the paper, we explore a counterfactual scenario examining the spread of ideas after removing Jesuit scholars from the network. Here, we present key statistics on the position and connectivity of Jesuits within the affiliation network. The network metric in Figure 3.22d quantifies how well-connected Jesuit-affiliated nodes are to the rest of the network. In our network structure, conductance reflects the extent to which Jesuits were integrated into mixed institutions rather than remaining isolated. The initially high conductance suggests that early Jesuits were present in diverse academic environments. Over time, the decline in conductance aligns with their increasing concentration within Jesuit institutions. Figure 3.23b shows the Homophily Index of Jesuits, by field between 1556 and 1767. Notably, the decreasing trend in inbreeding homophily among Jesuit scientists is likely a consequence of the rise of academies, which were largely centered on scientific disciplines and saw a relatively active participation from Jesuits.



(A) (B) (C) (D)

FIGURE 3.22: (a) Number of Jesuit scholars active in the network, 1556–1767. (b) Comparison of number of scholars by type: all (solid line), Jesuits (dotted line), and non-Jesuit scholars (dashed line), 1556–1767. (c) Fraction of Jesuit scholars active in the network, 1556–1767. (d) Conductance of Jesuits in the affiliation network, 1556–1767.

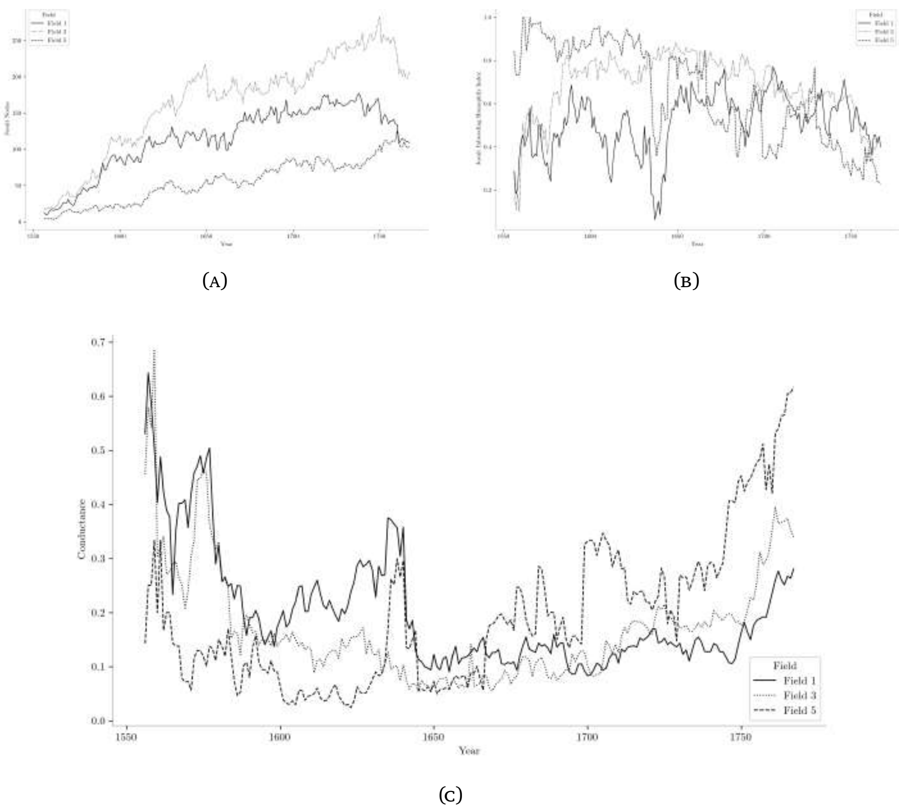


FIGURE 3.23: (a) Number of Jesuit scholars active in the network by field, 1556–1767. Field 1 stands for theology, field 3 for humanities, and field 5 for sciences. (b) Homophily Index of Jesuits by field, 1556–1767. (c) Conductance of Jesuits in the affiliation network by field, 1556–1767.